

1 CTGCAGCCAA CTTTGTGAC CATCTCCGCA ATGCCTTGG A CGTCCTGCAT
 51 AGAGAGCTTT TCCCCTTAGGT GCCCAGAGTC CTGGTCAACC TCGTGGACTT
 101 CCTGAACCCC ACTATCATGC GGCAGGTGTT CCTGGGAAAC CCAGACAAGT
 151 GCCCAGTGCA GCAGGCCAGA GCAGCATGCG CGAGCTGGTG GGGTCAGGCC
 201 GCTATGACAC GCAGGAGGAC TTCTCTGTGG TGCTGCAGCC CTTCTTCCAG
 251 AACATCCAGC TCCCTGTCTT GGCGCTTGA CCACTTGGAA GCAAAACAGA
 301 GACCCCTGGAC CTGAGAGCAG AGATGCCCAT CACCTGTCCC ACTCAGAATG
 351 AGCCCTTCCCT GAGAACCCCT CGGAATAGTA ACTACACGTA CCCCATCAAG
 401 CCAGCCATTG AGAACTGGGG CAGTGACTTC CTGTGTACAG AGTGGAAAGGC
 451 TTCCAATAGT GTTCCAACCT CTGTCCACCA GCTCCGACCA GCAGACATCA
 501 AAGTGGTGGC CGCCCTGGGT GACTCTCTGA CTACAGCAGT GGGAGCTCGA
 551 CCAAACAAC CCAGTGACCT ACCCACATCT TGGAGGGGAC TCTCTTGGAG
 601 CATTGGAGGG GATGGGAACT TGGAGACTCA CACCACACTG CCCAACATTC
 651 TGAAGAAGTT CAACCCCTAC CTCCCTTGGCT TCTCTACCAG CACCTGGGAG
 701 GGGACAGCAG GACTAAATGT GGCAGCGGA GGGGCCAGAG CTAGGGACAT
 751 GCCAGCCCCAG GCCTGGGACC TGGTAGAGCG AATGAAAAAC AGCCCCGACA
 801 TCAACCTGGA GAAAGACTGG AAGCTGGTC CACTCTTCAT TGGGGTCAAC
 851 GACTTGTGTC ATTACTGTGA GAATCGGAG GCCCACCTGG CCACGGAATA
 901 TGTTCAGCAC ATCCAACAGG CCCTGGACAT CCTCTCTGAG GAGCTCCCAA
 951 GGGCTTCGTT CAACGTGGTG GAGGTGATGG AGCTGGCTAG CCTGTACCCAG
 1001 GGCCAAGGCG GGAATGTGC CATGCTGGCA GCTCAGAACAA ACTGCACTTG
 1051 CCTCAGACAC TCGCAAAGCT CCCTGGAGAA GCAAGAACTG AAGAAAGTGA
 1101 ACTGGAACCT CCAGCATGGC ATCTCCAGTT TCTCCTACTG GCACCAATAC
 1151 ACACAGCGTG AGGACTTTGC GTTGTGGTG CAGCCTTTCT TCCAAAACAC
 1201 ACTCACCACCA CTGAACGAGA GAGGGGACAC TGACCTCACC TTCTCTCCG
 1251 AGGACTGTT TCACCTCTCA GACCGCGGGC ATGCCGAGAT GGCCATCGCA
 1301 CTCTGGAAACA ACATGCTGGGA ACCAGTGGGC CGCAAGACTA CCTCCAACAA
 1351 CTTCACCCAC AGCCGAGCCA AACTCAAGTG CCCCTCTCCT GAGAGCCCTT
 1401 ACCTCTACAC CCTGCGGAAC AGCCGATTGC TCCCAGACCA GGCTGAAGAA
 1451 GCCCCCGAGG TGCTCTACTG GGCTGTCCCA GTGGCAGCGG GAGTCGGCCT
 1501 TGTGGTGGGC ATCATCGGGG CAGTGGTCTG CAGGTGCAGG AGAGGTGGCC
 1551 GGAGGGAAAGA TCCTCCAATG AGCCTGCGCA CTGTGGCCCT CTAGGCCCCGG
 1601 GGGTGGGTCC TCACCCCTAAA CTCCCTATAG CCACTCTCTT CACCGCCCTC
 1651 TGCCCCAGCC ACTCCCGGCC ACCAGGACAT GCTTCAATGC CTGGTGCCAT
 1701 AGGAAGCCCA GGGGACAGTC ACAACTTCTT GGGGCTTGGG CTTCTTCCAG
 1751 GCCTATGCTC CTGAAATGGA TACATTTAAA TAAAGTCCAA AGCTATTTTA
 1801 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA

FEATURES:

5'UTR: 1 - 175
 Start Codon: 176
 Stop Codon: 1592
 3'UTR: 1595

Homologous proteins:

Top BLAST Hits

		Score	E
CRA 18000004885276	/altid=gi 464376 /def=sp Q05017 PHLX_RABIT P...	750	0.0
CRA 18000005150386	/altid=gi 3172337 /def=gb AAC40129.1 (AF045...	682	0.0
CRA 18000005121266	/altid=gi 2696236 /def=dbj BAA23813.1 (D636...	665	0.0
CRA 18000005181876	/altid=gi 7498717 /def=pir T20655 hypothetical...	228	9e-59
CRA 87000001028586	/altid=gi 7332170 /def=gb AAF60857.1 (AC024...	210	3e-53
CRA 18000005040393	/altid=gi 7508802 /def=pir T26083 hypothetical...	205	7e-52
CRA 89000000196200	/altid=gi 7293699 /def=gb AAF49069.1 (AE003...	200	3e-50
CRA 89000000199135	/altid=gi 7297015 /def=gb AAF52285.1 (AE003...	189	6e-47
CRA 18000004979533	/altid=gi 7499049 /def=pir T16060 hypothetical...	161	2e-38
CRA 18000005184633	/altid=gi 7506410 /def=pir T24016 hypothetical...	152	9e-36
CRA 18000005184632	/altid=gi 7506411 /def=pir T24015 hypothetical...	122	1e-26
CRA 18000005182912	/altid=gi 7500588 /def=pir T21835 hypothetical...	119	6e-26
CRA 87000001028649	/altid=gi 7332235 /def=gb AAF60922.1 (AC006...	111	2e-23

BLAST dbEST hits:

gi 2079883 /dataset=dbest /taxon=9606 ...	724	0.0
gi 11593761 /dataset=dbest /taxon=960... .	670	0.0
gi 7037501 /dataset=dbest /taxon=9606... .	654	0.0
gi 12241943 /dataset=dbest /taxon=96... .	632	e-179
gi 10367787 /dataset=dbest /taxon=960... .	575	e-161
gi 9969781 /dataset=dbest /taxon=960... .	547	e-153
gi 7667765 /dataset=dbest /taxon=9606... .	531	e-148
gi 12241345 /dataset=dbest /taxon=96... .	519	e-145
gi 2080047 /dataset=dbest /taxon=9606	468	e-129
gi 7946640 /dataset=dbest /taxon=960... .	323	7e-86

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

gi|2079883 Mixed (melaonocyte, fetal heart, pregnant uterus)
gi|11593761 Kidney
gi|7037501 Whole blood
gi|12241943 Lung, normal
gi|10367787 brain glioblastoma
gi|9969781 Prostate
gi|7667765 Colon
gi|12241345 Lung, normal
gi|2080047 Mixed (melaonocyte, fetal heart, pregnant uterus)
gi|7946640 Colon

Expression information from PCR-based tissue screening panels:

Leukocyte

1 MRELVGSGRY DTQEDFSVVL QPFFQNIQLP VLALEPLGSK TETLDLRAEM
51 PITCPTQNEP FLRTPRNSNY TYPPIKPAIEN WGSDFLCTEW KASNSVPTSV
101 HQLRPADIKV VAALGDSLTT AVGARPNNSS DLPTSWRGLS WSIGGDGNLE
151 THHTTLPNLIK KFNPYLLGFS TSTWEGTAGL NVAAEGARAR DMPAQAWDLV
201 ERMKNSPDIN LEKDWKLVTL FIGVNNDLCHY CENPEAHLAT EYVQHIIQQAL
251 DILSEELPRA FVNVEVMEI ASLYQGQGGK CAMLAAQNNC TCLRHSQSSL
301 EKQELKKVNW NLQHGISSFS YWHQYTQRED FAVVVQPFFQ NTLTPLNERG
351 DTDLTFFSED CFHFSDRGHA EMAIALWNMM LEPVGRKTTS NNFTHSRAKL
401 KCPSPESPYL YTLLRNSRLLP DQAEEAPEVL YWAVPVAAGV GLVVGIIIGTV
451 VWRCCRGGRR EDPPMSLRTV AL

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 5

1	69-72	NYTY
2	127-130	NNSS
3	128-131	NSSD
4	289-292	NCTC
5	392-395	NFTH

[2] PDOC00004 PS00004 cAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

386-389 RKTT

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 7

1	7-9	SGR
2	64-66	TPR
3	135-137	SWR
4	326-328	TQR
5	365-367	SDR
6	412-414	TLR
7	466-468	SLR

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 6

1	12-15	TQED
2	39-42	SKTE
3	56-59	TQNE
4	172-175	STWE
5	298-301	SSLE
6	326-329	TQRE

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 8

1	123-128	GARPNN
2	138-143	GLWSI
3	144-149	GGDGNL
4	147-152	GNLETH
5	179-184	GLNVAA
6	278-283	GGKCAM
7	441-446	GLVVGI
8	445-450	GIIGTV

[6] PDOC00009 PS00009 AMIDATION
Amidation site

Number of matches: 2

1	384-387	VGRK
2	457-460	GGRR

[7] PDOC00016 PS00016 RGD
Cell attachment sequence

349-351 RGD

[8] PDOC00200 PS00228 TUBULIN_B_AUTOREG
Tubulin-beta mRNA autoregulation signal

1-4 MREL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	164	184	0.694	Putative
2	432	452	1.956	Certain

BLAST Alignment to Top Hit:

>CRA|18000004885276 /altid=gi|464376 /def=sp|Q05017|PHLX_RABIT
PHOSPHOLIPASE ADRAB-B PRECURSOR /dataset=nraa
/length=1458
Length = 1458

Score = 750 bits (1915), Expect = 0.0
Identities = 368/502 (73%), Positives = 407/502 (80%), Gaps = 33/502 (6%)

Query: 1 MRELVGSGRYDTQEDFSVVLQPFFQNIQLPVLA----- 33
+RELV SGRYDT+EDFSVVLQPFF +IQLPVL
Sbjct: 955 LRELVESGRYDTREDFSVVLQPFFHSIQLPVLQDGRLDTSFFAPDCVHPNQKFHSQLSRA 1014

Query: 34 -----LEPLGSKTETLDLRAEMPTCPTQNEPFLRTPRNSNYTYPPIKPAIENWGSDFLCT 88
LEPLG KT+ LDL A + +TCPTQNEPFLRT RNS+YTYP +PA+ENWGSDFLCT
Sbjct: 1015 LWRNMLEPLGGKTDALDLTAAITLTCPTQNEPFLRTFRNSDYTYSRPAVENWGSDFLCT 1074

Query: 89 EWKASNSVPSTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSWSIGGDGN 148
W AS VP SVH+L+P DIKVVAALGDSL T A+GARP+NSSD P WRGLSWSIGGDG
Sbjct: 1075 AWNASRGVPNSVHELQPGDIKVVAALGDSLTLAMCARPSNSDPPMFWRGLSWSIGGDGA 1134

Query: 149 LETHTTLPNILKKFNPYLLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLVERMKNSPD 208
LETHTTLPNILKKFNP +LGFST T EGT GLNVA +GARA+DMPAQA DLVERM+NSP+
Sbjct: 1135 LETHTTLPNILKKFNPSILGFSTGTLEGTMGLNVAVQGARAQDMPAQARDLVERMRNSPE 1194

Query: 209 INLEKDWKLVTLFFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVEVM 268
I+LEKDWKLVTLF+G NDLCH+CENPE EYVQHIQQALD+L EELPR FVNVEVM
Sbjct: 1195 IDLEKDWKLVTLFVGGNNDLCHFCENPEGSSEGEYVQHIQQALDVLYEELPRTFVNVEVM 1254

Query: 269 ELASLYQGQGGKCA-MLAAQNNTCLRHSQSSELEKQELKKVWNWLQHGIISSFSYWHQYTQ 327
ELA L+Q QGG+CA +LAAQ++CTC ++SQSS+E QELKKVWNWLQ G+S SY HQY Q
Sbjct: 1255 ELAGLHQDQGGRCATLLAAQSHCTCFKYSQSSVEMQELKKVWNWLQSGLSRLSYSHQYVQ 1314

Query: 328 REDFAVVVQPFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEAMAIALWNMLEPVGRK 387
REDFAVVVQPFQNTL PLN RGDTDLTFFS+DCFHFS+RGHAEMAIALWNMLEPVG K
Sbjct: 1315 REDFAVVVQPFQNTLVLNGRGDTDLTFFSDDCFHFSERGHAEAMAIALWNMLEPVGHK 1374

Query: 388 TTSNNFTHSRAKLKCPSPESPYLYTLRNSRLLPDQAEAAPEVLYWAVPVAAGVGLVVGII 447
TTSNNFT+SR KLKCPS+SPYLYTLRNSRLLPDQAE P VLYWAVPVAAG GL++GI+
Sbjct: 1375 TTSNNFTYSRTKLKCPSPDSPYLYTLRNSRLLPDQAEADPTVLYWAVPVAAGAGLLIGIL 1434

Query: 448 GTVVWRCRRGGRRDPMSLRT 469
V R R REDPP+SL T
Sbjct: 1435 AMVAGRGMCRPREDPPLSLST 1456

Score = 334 bits (847), Expect = 2e-90
Identities = 191/445 (42%), Positives = 251/445 (55%), Gaps = 33/445 (7%)

Query: 4 LVGSGRYDTQEDFSVVLQPFFQNIQLPVLA-----LEPLGSKTETL 44
L+ S +Y+TQE F+VV QPFF L L +EP+G K E
Sbjct: 264 LLASSKYNTQESFAVVFPQFFYESSLSALLAEPPLQDPTTLALSLWNRMMEPIGRKEEF 323

Query: 45 DLRAEMPTCPTQNEPFLRTPRNSNY----TYPPIKPAIENWGSDFLCTEWKASNSVPSTV 100
+ P+ CPTQ P+L T RNS + P G++ C + S+SVPTSV
Sbjct: 324 SEKERKPLRCPTQESPYLYFTYRNSGQLTRVSQPQGKLEVREGTEIRCPDKDPSDSVPTSV 383

Query: 101 HQLRPADIKVVAALGDSLTTAVGA--RPNNSSDLPTSWRGLSWSIGGDGNLETHTLPNI 158
H+L+PADIKV+ A+GDSL T GA +P N D+ T +RGLSWS+GGD N+ T TTL NI
Sbjct: 384 HRLKPADIKVIGAMGDSLTAGNGAGSQPGNILDVLTQYRGLSWSVGGDQNISTVTTLANI 443

Query: 159 LKKFNPYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMKNSPDINLEKDW 215
L++FNP L GFS T T A N A GARA + QA LV MKN IN ++DW

Sbjct: 444 LREFNPSLQGFSVGTGRETTSQAFFNQAVAGARADGLIPQAQRLVALMKNDTRINFOEDW 503

Query: 216 KLVTLFIGVNDLCHYCENPEAHLATEYVQHQIQQALDILSEELPRAFVNVEVME---LAS 272
K++T+FIG NDLC +C +P + + +I ALDIL E+PRAFVN+V+V+E L

Sbjct: 504 KIITVFIGGNNDLCDFCNDPVRYSPQNFTDNIGTALDILHAEIPRAFVLVKVLEISKLRE 563

Query: 273 LYQQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFSYWHQYTQREDF 331
LYQ C + ++ C C L+ +S E L+ Q +Y R+DF

Sbjct: 564 LYQETKVSCPRMILRSLCPVLFDDNSTEIASLIETIKEYQERTQQLIDSGRYDTRDDF 623

Query: 332 AVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSRGRHAEMAIALWNNMLEPVGRKTTSN 391
VV+QPFF+ P + G D +FF+ DCFHFS + HA A ALWNNMLEPVG+KTT N

Sbjct: 624 TVVLQPFFEKFVNMPKTQDGLPDNSFFAPDCFHFSKAHAHAASALWNNMLEPVGQKTTHN 683

Query: 392 NFTHSRAKLKCPSPEPYLYTLRNS 416
+F + CP+ P+L T +NS

Sbjct: 684 DF-EGAVNITCPNQVWPFLSTYKNS 707

Score = 323 bits (819), Expect = 3e-87

Identities = 181/456 (39%), Positives = 261/456 (56%), Gaps = 51/456 (11%)

Query: 2 RELVGSGRYDTQEDFSVVLQPFFQNIQLPVLA----- 33
++L+ SGRYDT++DF+VVLQPFF+ +P

Sbjct: 609 QQLIDSGRYDTRDDFTVVLQPFFEKFVNMPKTQDGLPDNSFFAPDCFHFSKAHAHAASAL 668

Query: 34 -----LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTYPPIKPAIENWGSDFLCTE 89
LEP+G KT D + ITCP Q PFL T +NS ++ +G+ C +

Sbjct: 669 WNNMLEPVGQKTTHNDFEGAVNITCPNQVWPFLSTYKNS-----VQGFGTWLPCRD 719

Query: 90 WKASNSVPSTVHQLRPADIKVVAALGDSLTTAVG--ARPNNSSDLPTSWRGLSWSIGGDG 147
S S PTSVH LRPADI+VVAALGDSL +G ++PN+ SD T +RGLS+S GGDG

Sbjct: 720 RSPSASPPTSVHALRPADIQVVAALGDSLTAGIGIGSKPNLSDGTTQYRGLSYSSGGDG 779

Query: 148 NLEHTTLPNILKKFNPYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMK 204
+L+ TTLPNIL++FN L+GF+ T + + A N A GA+ARD+ +Q LV+RMK

Sbjct: 780 SLNDVTTLPNILRQFNSNLMGFAVGTGDASGTNAFFNQAVPGAKARDLMSQVQLTVQRMK 839

Query: 205 NSPDINLEKDWKLVTLEFIGVNDLCHYCENPEAHLATEYVQHQIQQALDILSEELPRAFVN 264
+ +N ++DWK++T+ IG +DLC YC + + A + H++ ALD L E+PRA VN+

Sbjct: 840 DDHRVNFQEDWKVITVQIGASDLCDYCTDSNLYSAANFYDHLDALDALHREVPRALVNL 899

Query: 265 VEVME---LASLYQGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFS 320
V+ M ++ G KC + A C C L ++S E L+ + Q +

Sbjct: 900 VDFMNPSVTQVFLGNPDKCPVQQASALCNCVLSPRENSYELARLEALAQAYQSSLRELV 959

Query: 321 YWHQYTQREDFAVVVPQFFQNTLTPLNERGDTDLTFFSEDCFHFSRGRHAEMAIALWNNM 380
+Y REDF+VV+QPFF + P+ + G D +FF+ DC H + + H++++ ALW NM

Sbjct: 960 ESGRYDTREDFSVVLQPFFHSIQLPVLQDGRLDTSFFAPDCVHPNQKFSQLSRALWRNM 1019

Query: 381 LEPVGRKTTSNFTHSRAKLKCPSPEPYLYTLRNS 416
LEP+G KT + + T + L CP+ P+L T RNS

Sbjct: 1020 LEPLGGKTDALDLT-AAITLTCPTQNEPFLRTFRNS 1054

Score = 137 bits (341), Expect = 3e-31

Identities = 107/338 (31%), Positives = 161/338 (46%), Gaps = 42/338 (12%)

Query: 85 FLCTEWKASNSVPT-SVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSWSI 143
F C + SVP+ SVH LRP+DIK VAA+G+ T + + T R +

Sbjct: 46 FPCDPKTLAESVPSSESVHSLRPSDIKFVAAIGNVETAPDSGADDLEEQDGTEKRPEQACM 105

Query: 144 GGDGNLEHTTLPNILKKFNPY-LLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLVER 202

G T L +I+ +F+P L+ T + G A D+ QA +LV
 Sbjct: 106 G-----VVTVLSDIIGRFSPSALMPLCPET-----RLVPRGG-AEDLWMOATELVR 151

Query: 203 MKNSPDINLEKDWKLVTLFIGVNDLCHYCENPEAH-LATEYVQHQALDILSEELPRAF 261
 M+ +P ++ E DWKL+ +F C C + + L + + + LD L +E+P+AF

Sbjct: 152 MRENPKQLDFEHDWKLINVFFSNTSQCFPCPSAQKGLVLGGMMDKLTLDYLQQEVPKAF 211

Query: 262 VNVVEVMELASLYQGQGGKCAMLAAQNNCTCLRHSQSSLEKQELKKV--NWNLQHGISSF 319
 VN+V++ ELA+ + + G + A C CLR E +L KV W+ S

Sbjct: 212 VNLVDLSELAAFSRWRQG-AQLSPAEEPCRCLR-----ETSQLTKVLTQWSYLEAWDSL 264

Query: 320 SYWHQYTQREDFAVVVQPFF-QNTLTPNLERGDTDLTFFSEDCFHFSDRGHAEMAIALWN 378
 +Y +E FAVV QPFF +++L+ L + +A++LWN

Sbjct: 265 LASSKYNTQESFAVVVFQPFYESSLSALLAEPL-----QDPTTLALSILWN 310

Query: 379 NMLEPVGRKTTSNNFTHSRAKLKCPSPEPYLYTLRNS 416
 M+EP+GRK + R L+CP+ ESPYLY+T RNS

Sbjct: 311 RMMEPIGRKEEPFS-EKERKPLRCPTQESPYLYTYSRNS 347

>CRA|18000005150386 /altid=gi|3172337 /def=gb|AAC40129.1| (AF045454)
 phospholipase B [Cavia porcellus] /org=Cavia porcellus
 /taxon=10141 /dataset=nraa /length=1463
 Length = 1463

Score = 682 bits (1741), Expect = 0.0
 Identities = 348/505 (68%), Positives = 389/505 (76%), Gaps = 38/505 (7%)

Query: 1 MRELVGSGRYDTQEDFSVVLQPFQNIQLPVLA----- 33
 MRELV SGRYDT+EDFSVVLQPF NI+LP+L

Sbjct: 954 MRELVESGRYDTREDFSVVLQPFILNIRLFILEDGRPDTSFFAPDCINPGQKFHSQLSRA 1013

Query: 34 -----LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTYPKPAIENWGSDFLCT 88
 LEP+GSKT+TLDL A++ + CPTQ EPFLRTP+NS+YTYP KPAIENWGSDFLCT

Sbjct: 1014 LWVNMLEPGSKTDTLTADISLPCPTQEEPFLRTPQNSDYTYPTKPAIENWGSDFLCT 1073

Query: 89 EWKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDLPTSWRGLSWSIGGDGN 148
 EWK SNSVPTSVH+L+PADIKVVAALGDSLTTAVGAR +NSSDL SWRGLSWSIGGDG

Sbjct: 1074 EWKPSNSVPTSVHKLQPADIKVVAALGDSLTTAVGARASNSSLMSWRGLSWSIGGDGA 1133

Query: 149 LETHTTLPNILKKFNPyLLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLVERMKNSPD 208
 LETHTTLPNILKKFN + GFST T E TAG NVA E ARARDMPQAQ DLVERMK S +

Sbjct: 1134 LETHTTLPNILKKFNPSIFGFSTGTLEETAGFNVAEEARARDMPAQARDLVERMKASTE 1193

Query: 209 INLEKDWKLVTLFIGVNDLCHYCENPEAH-LATEYVQHQALDILSEELPRAFVNVE-V 267
 INLE DWKL+TLFIG NDLCHYC+NPE H A EYVQHQALDILSEELPRAF+NVV+ +

Sbjct: 1194 INLEMDWKLITLFIGSNDLCHYCDNPENHSAEEYVQHQALDILYELPRAFINVVDII 1253

Query: 268 MELASLYQGQGGKC-AMLAAQNNCTCLRHSQSSLEKQELKVNWLQHGISSFYWHQYT 326
 MELA L+QGQGG C A+L AQ+ C+CLRH SS QELKKV WNLQ +S SY +YT

Sbjct: 1254 MELAGLHQGQGHCTALLPAQSTCSCLRHFPSSPVQELKVTWNLQSDMSRLSYQEKYT 1313

Query: 327 QREDFAVVVQPFFQNTLTPNLERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNMLEPVGR 386
 QREDFAVVVQPFFQNTL PL++ G TD TFFSEDC HFS+RGHAEMAIALWNNMLEPVG

Sbjct: 1314 QREDFAVVVQPFFQNTLIPLDKLGSTDPTFFSEDCLFHFSERGHAEMAIALWNNMLEPVGH 1373

Query: 387 KTTSNFTHSRAKLKCPSPEPYLYTLRNSRLLPDQAEAAPEVLYWAVPVAAG---VGL 442
 KTT NNFT++R KLKCPS ESPYLYTL+NS LP Q E+A V V AA VGL

Sbjct: 1374 KTTFNNFTYNRTKLKCPSTEPYLYTLQNSLSPVQTEKASGVAPGIVSAAAAGGLLVGL 1433

Query: 443 VVGIIGTVVWRCCRGGRRDPMSL 467
 +VGI+ +W R +++ PP S+

Sbjct: 1434 IVGILAVSLWSSFRQQKSPPEV 1458

Score = 348 bits (884), Expect = 7e-95
 Identities = 199/442 (45%), Positives = 257/442 (58%), Gaps = 31/442 (7%)
 Query: 4 LVGSGRYDTQEDFSVVLQPFFQNQJQLPV-----LAL-----EPLGSKTETLDL 46
 L+ S ++ QE F+VV QPFF + PV LAL +P+G K E
 Sbjct: 265 LLASSSFNDQESFAVVFPQFFYEVSSPVEEPPSQDPTTLALSLWNNMMKPVGQKDEPFST 324
 Query: 47 RAEMPITCPTQNEPFLRTPRNSNYTPI---KPAIENWGSDFLCTEWKASNSVPTSVHQ 102
 P+ CP+Q P+L T RNSNY + + E G++ C + S+S PTSVH+
 Sbjct: 325 IERRPMKCPSQESPYLFTYRNSNYQSRLLKRQRQHKEREGTEIRCPDKDPSDSTPTSVHR 384
 Query: 103 LRPADIKVVAALGDSLTTAVGA--RPNNSSDLPTSWRGLSWSIGGDGNLEHTTLPNILK 160
 L+PADIKV+ ALGDSL T GA RP N D+ T +RGLSWSIG D N+ + TTLPNIL+
 Sbjct: 385 LKPADIKVIGALGDSLTAGNGAGSRPGNILDVLTEYRGLSWSIGADHNISSVTTLPNILR 444
 Query: 161 KFNPYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMKNSPDINLEKDWL 217
 +FNP L GFST T + A N A GARA D+ QA LV+ MKN IN E+DWK+
 Sbjct: 445 EFNPSLKGFGSTGTGKANSVGAFFNQAVAGARAGDLIPQARTLVDMKNHTSINFEEDWKI 504
 Query: 218 VTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVEVME---LASLY 274
 +T+FIG NDLC +C +P + + +I+QALDIL E+PRAFVN+V+V++ L LY
 Sbjct: 505 ITVFIGGNDLCDFCSDPVTNSPENFTDNIRQALDILHAEVPRAFVNVMVKVLQIVNLRELY 564
 Query: 275 QQQGGKCAMLAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFSYWHQYTQREDFAV 333
 + C L +N C C L +S E + L +N Q +Y REDF V
 Sbjct: 565 KDSRVSCPRLLRLNCRCVLLPDDNSTELESIDINKKYQERTHQLIESGRYDTREDFTV 624
 Query: 334 VVQPFQNTLTPNLERGDTDLTFFSEDCFHFSRGRHAEMAIALWNNMLEPVGRKTTSNF 393
 V+QPF+ P G D T F+ DCFHFS + HA A ALW NMLEPVG+KTT NNF
 Sbjct: 625 VLQPFFEKEVDIPKTSEGIPDNTSFAPDCFHSSKTHARAASALWKNMLEPVGQKTTQNNF 684
 Query: 394 THSRAKLKCSPSPESPYLYTLRN 415
 +S + CP+ PYL T +N
 Sbjct: 685 ENS-IDIICPNQAFPYLSTYKN 705

Score = 314 bits (795), Expect = 2e-84
 Identities = 178/455 (39%), Positives = 259/455 (56%), Gaps = 51/455 (11%)
 Query: 3 ELVGSGRYDTQEDFSVVLQPFFQNQJQLPVLA----- 33
 +L+ SGRYDT+EDF+VVLQPFF+ + +P +
 Sbjct: 609 QLIESGRYDTREDFTVVLQPFFEKEVDIPKTSEGIPDNTSFAPDCFHSSKTHARAASALW 668
 Query: 34 ---LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTPIKPAIENWGSDFLCTEW 90
 LEP+G KT + + I CP Q P+L T +N IE G+ C E
 Sbjct: 669 KNMLEPVGQKTTQNNFENSIDIICPNQAFPYLSTYKNG-----IEGHGTWLTCRER 719
 Query: 91 KASNSVPTSVHQLRPADIKVVAALGDSL--TAVGARPNNSSDLPTSWRGLSWSIGGDGN 148
 S S PTSVH LRPAD++VVAALGDSL + +G++P + +D+ T +RGLS+S GGDG+
 Sbjct: 720 TPSASPPTSVHALRPADVRVVAALGDSLTAGSGIGSKPGDLADVITQYRGLSYSSGGDGS 779
 Query: 149 LETHTTLPNILKKFNPYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMKN 205
 L TTLPNIL++FN L G++ T + + A LN A GA+A ++ +Q LV++MK+
 Sbjct: 780 LMNVTTLPNILREFNSNLTYAVGTGDA5NTNAFLNQAVPGAKAEELMSQVKTLVQKMKD 839
 Query: 206 SPDINLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVV 265
 P IN +DWK++T+ IG NDLC++C + + + + H+ ALDIL E+PRA VN+V

Sbjct: 840 DPRINFHEDWKVITVLIGTNDCNHCTDLDLYSSANFFNHLLNALDILHREVPRALVNLV 899
 Query: 266 EVME---LASLYQGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFSY 321
 + M + ++ G KC + A C C L ++S E + + Q +
 Sbjct: 900 DFMNPSIMRQVFLGNPDKCPVQQASILCNCVSLRENSYELARMDALTRAYQSSMRELVE 959
 Query: 322 WHQYTQREDFAVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNML 381
 +Y REDF+VV+QPFF N P+ E G D +FF+ DC + + H++++ ALW NML
 Sbjct: 960 SGRYDTREDFSVVLQPFFLNIRLPILEDGRPDTSSFPDCINPGQKFHSQLSRALWVNML 1019
 Query: 382 EPVGRKTTSNNFTHSRAKLKCPSPEPYLYTLRNS 416
 EPVG KT + + T + L CP+ E P+L T +NS
 Sbjct: 1020 EPVGSKTDTLDLT-ADISLPCPTQEEPFLRTPQNS 1053

Score = 155 bits (389), Expect = 8e-37
 Identities = 110/351 (31%), Positives = 166/351 (46%), Gaps = 48/351 (13%)

Query: 85 FLCTEWKASNSVPT-SVHQLRPADIKVVAALGDSLTTAVGARPNNSSDPLTS---WRGLS 140
 F C+ K ++P+ SVH L PADIK++AA+GD T N + T WRG
 Sbjct: 45 FSCSPKKLGLNMPSESVHTLTPADIKLIAAIGDMETPPDSGAVNLDTSERTEKEPWRCM 104
 Query: 141 WSIGGDGNLEHTTLPNILKKFNPYLLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLV 200
 + T L +I+ FNP +L + W AA ++ QA +LV
 Sbjct: 105 GMM-----TVLSDIISHFNPSVLLPTCPPWRS-----AAVRGGVEELRTQAEELV 149
 Query: 201 ERMKNSPDINLEKDWKLVTFIGVNDLCHYCENEAAHLATEYVQHIQQ---ALDILSEEL 257
 +K +P ++ ++DWKL+ +F LC+ C P AH + ++ + L L +E+
 Sbjct: 150 SSLKKNPQLDFQQDWKLINVFFSNASLCYLC--PSAHENGPLMSNMDKLAGILHYLHQEV 207
 Query: 258 PRAFVNVEVMELASLYQGQGGKCAMLAAQNNCTCLRHSQSSLEKQELKKVNWNLQHGIS 317
 PRAFVN+V++ E+ ++ + G + C C K + + W+ Q
 Sbjct: 208 PRAFVNLDLFEVVAMPRWHQGTMRLSRPSPEACGC----SGETSKLDTVVMQWSYQETWD 263
 Query: 318 SFSYWHQYTQREDFAVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALW 377
 S + +E FAVV QPFF +P+ E D T +A++LW
 Sbjct: 264 SLLASSSFNDQESFAVVFQPFYEVSSPVEEPPSQDPT-----TLALSLW 308
 Query: 378 NNMLEPVGRKTTSNNFTHSRAKLKCPSPEPYLYTLRN----SRLLPDQAE 424
 NNM++PVG+K + T R +KCPS ESPYL+T RN SRLL Q +
 Sbjct: 309 NNMMKPGQKDEPFS-TIERRPMKCPSEQSPYLFTYRNSNYQSRLLKRQRQ 358

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00657	Lipase/Acylhydrolase with GD\$-like motif	158.4	1.6e-45	1
PF01347	Lipoprotein amino terminal region	1.9	6.1	1
CE00543	CE00543 steroid_receptor_N10	-0.8	1.5	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01347	1/1	155	172 ..	636	653 ..	1.9	6.1
PF00657	1/1	110	233 ..	1	146 []	158.4	1.6e-45
CE00543	1/1	230	254 ..	456	480 ..	-0.8	1.5

1 ATTCTGCAGC CAACTTTGTT GACCATCTCC GCAATGCCCT GGACGTCTG
51 CATAGAGAGG TGGGTGGGG GCTTCCACAA GCTGGTAACA GCTCAAGCAT
101 GGTGAGGGTG AAGGTGGATG GGGGGAAAGA ATGAGAGAAG AACCCCTTC
151 TCTCAAGGAG ACAGCCAAGG GCATGGANNN NNNNNNNNNN NNNNNNNNN
201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NGTGCATTG
951 CCGCTGCAGC CCCTTGGGT GGCACCCATG GAGTTGTGTG ATGTACGGCC
1001 TGTAAGGCCT TACGAGGTAG CCCTGTATAG ACTCCTCCCC AGAACTCAAC
1051 TCCAGAAAAGA CCAAGCTGGA TTGCTAAAGG AACCCATTCC TAGGGGCCCT
1101 GAGACAGCCC CAGGAAGAAG TGCCTGGAGC CCCCTCTCA TCTGCAGCTT
1151 TTCCCTTAGG TGCCCAAGGT CTCGGTCAAC CTCGGACT TCCCTGAACCC
1201 CACTATCATG CGGCAGGTGT TCCTGGGAAA CCCAGACAAG TGCCCAGTGC
1251 AGCAGGCCAG GTAGGCAGGT CCTGGCTGTC CCCACACTGG AGATGCCCTC
1301 ACCTCCTGGT CTGGCCCCACA TGCAGTGGTG ATGCCCTCAGG GTCTTTGTGA
1351 CTTGGCTAT CCATGTGTCC AAGTCTGTAA AGGAGGACTT CTGCCAGAAC
1401 GTCCCCTTCC AGAGGCTGGA GCCATGACTC CCCCTGTTACC CAACTTCAAG
1451 GTGCCTGGCA GGAACCTCTA TGATACCAGG CAGGCCACAGA GGGGAGGGAT
1501 CAAAGTTGGG ACAGAGGCTG GTGTTTGAGA GACAGGATAG CCTAGACTGT
1551 GAACATGGGC AGTGGTTAGG GATGTAGACA TATGTGGTCA AACTGTAACA
1601 GAAAGCAAGG AAAAGGTACA AGCAACTCAG TTACCTTTAG GGGAAAGAAGA
1651 GAATTAGGAG GGACACAGGG AGCTTCAAAC TGGGAGTGTGTT TTGTTTCTTA
1701 AACTGGGCCA TAAGTACATG GATGTGTGTT TTATTATTCT TTATATCTTA
1751 CACATCTATT TACTCAGCA ATCTTACAGA ACTTCTGTG TACCAAGGCAT
1801 TGTTTCAAGT GCTTAGAAA TCTCTCTCTT AAGTAGATGT GATGGGTGTG
1851 AAATAATICA TGATGAAACC AAAGGGGACA CAGTAGGGCA CTCATGTGAA
1901 AGAAGGAGAG GTCTAAGGCA TAGCATCAGA GGCCCCAAAA TATCAGCTCC
1951 AACACCAGAG GATGCATTTC CTTTTTAATT AAACACTAAA TTTTCACTGC
2001 CCAAATTCAAT TTGCTCAGCT GAATAATCGG TTGCAGGCC AGCACCTGCA
2051 GTCCAACACT TGTGCTCTGT TGGTATGAGA GGGTGCCTCAT TCCCACGCTG
2101 GCTCCCTCCC TCAGGCCATC TCCAGTGCCTC CTGCCAGGCC TGAAGCCTGC
2151 CCCTGAGCAT GTGCGCCAGA GCCTCAAGGC TTGAGTGCTC CTAACCAGG
2201 GCGGGAGGGGA GCCTCTCCAC CCCTCCCCCTG AACCTGGGCA ATCAGAACCA
2251 GCCCCGTATG GAAGCCTGAG CTCTGGGGCC TCCTGCCTCC CCCTCTTTGT
2301 GCAGCGTTT GTGTAACTGC GTTCTGACCC TGCAGGAGAA CTCCCAAGAG
2351 CTAGCCAGGC TGGAGGCCTT CAGCCGAGCC TACCGGGTAA GACCAAGAAC
2401 GGCACCATGTC TGTGCTCTCT CCCCTACGTT CACTCTAACAA CACAGCCCCAG
2451 AGCCCCCTAGA GGAGGCACAC AGGGAAAGGAA AAGCTGGTCA GGGATTGTGG
2501 GGAGAGGGGG AGCAGCCTGG GTGCCTTCTCT GTGCTCAGC TGACTGTGGT
2551 GTCTCAGGTG CCCTGGTTGG AATCATCCCA GTAGGATCCA GGTGGAAAAG
2601 CCCTCATGGC CCAGCTACCG TTGAGGGCTT AACCCCAACT CCTGGCCCGT
2651 AGGCCCTGGAT GCCTCATGAG ACCACCTTTC CCTCCCCCAC TCCCACTCCA
2701 AAGGCAGGTG CCGAGCCTCT TCAGAGAGAG TAGTCTACTT ACAATCAAAA
2751 GGGACTTCTCT GCCTAGCCCT TCAGAGAGAG TAGTCTACTT ACAATCAAAA
2801 CAAAAAAGGTG ACCCAACCTG TTTCCAAATT CTCTGGAAAG GGACTTGCCTC
2851 TCAGGTGATT TGTGTTCTCA AGGGAAAGGC TGAGTCGGCC CCTCCATCCA
2901 GGGAGATGGA CTGCCACCA CCCCTACTCT TGCCTCACTG GGTCTGGGC
2951 CCACCCAGGG CCTGGGCTGA AGACCCCTGTG CATGTGTCCC CAGAGCAGCA
3001 TGCGCGAGCT GGTGGGGTCA GGCGCTATG ACACGCAGGA GGACTTCTCT
3051 GTGGTGCCTGC AGCCCTTCTT CCAGAACATC CAGCTCCCTG TCTGGTGGT
3101 ATGTCCCTG CCCTCGCCCA TGGTACTCTT TTAGAGGAAG AAATGCAAGG

FIGURE 3, page 1 of 33

3151 CAGAATTGCC AGTTGCTTCC ACGAGCATGT GCATAAAATG GGAAAGACAC
3201 AGCTCTCCAG ACGCTGNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3751 TTGTGGGATC GCTCTGATCT CTCTGTTAAG TGAATGGGCC CTGTGGTGGC
3801 TGGTGACCTG GAGCACTCCA GGGGAAGGAA GGTGTTGAGT GGTCACTCCC
3851 AGGGCCAAGT CCGCTGGTGG TGGCTCCCTC TGAACCAATA GGATCTTGAG
3901 GGGGTATATT GGTCTCTTC AGGATGGGCT CCCAGATACT TCCTTCTTTG
3951 CCCCCAGACTG CATCCACCCA AATCAGAAAT TCCACTCCC GCTGGCCAGA
4001 GCCCTTGGA CCAATATGGT AAAATAAGTG GGGTGTTCCT TGTTCCTCTGG
4051 GGTTCTAGTC TAGGGCAGGG CACCAGCCCC TATAGAATGG AGTCTTGCAA
4101 GTGAGGCCTGA GGGGGCAGTG GCTGGTACAT CTATAAACGT CTATGCAGTT
4151 GGAAATGCCG AGTCCTTAAG AGTCTGCTCA GCCTGGGCTC AACTGCACCC
4201 TCTCCTCAGA GCTTGAACG CTGAGGAGGG ACCTCTCTAC AGAAATGCAA
4251 GCCCAAACCC CCATATTCTAT TCCACTTTCC CTATGTGCCG GCCACCATGT
4301 TAGGCAGTTT AAGCCACGTT ATCTCAITTTA AGGCTCTGCA TATCCCTGCT
4351 AGAGAACCAT GACGAGCTC CAAGGAAGCT ACTCCCCAGAG AAGCAAAGCG
4401 ACTGGCCCAA AACCCCCACGG CTGGCAACTG GCAGAGGCCG AAGTGGGAGC
4451 CAAGCCCCCT GAAATCGAGT TCTGAGCTTT CCCCACTGCA GGATTCTGCC
4501 AGGGGAATGTT CACTTCCATG GAAACAAAACT ACTACACCCG TGTCTCTCTT
4551 TTCTTCCCTG ATCAGCTTGA ACCACTTGA AGCAAAACAG AGACCCCTGGA
4601 CCTGAGGACA GAGATGCCA TCACCTGTCC CACTCAGGTA GTAGGGGAGG
4651 ACCTGCCTGG CTCCCTCTCCA CAAACCAGGGCACACAGCTC GCCCTACCCA
4701 CTTCGTCCTC CACACAGCT TCCTCAGTAC CCATCTTGCC CCCTTACTGAA
4751 GGCCTGAGAG ATTGGAGAG TGGAGGGAG TCCATGAGGA TGACAGGGGG
4801 AGGTGAGAGG GGAGACAAGA GTGCAGCTGT CAATTGGAAAC AGGAGATGCA
4851 GCAGGGAGAG GAGGCCCTGGG CCCCAGCAGA GGGAGAGGAT CCCGGTGAGA
4901 AAAGTGGGCT CCTGAGAGAG GAAATCAGGA TGCCAGGAAA ATGGCAGGAG
4951 GGCTTCTCTT AGCAGTGGGT TTTGGGGCAG ATGAAAAAAAT CTGACTGCAG
5001 GTTAGAGGGC CCAGGCAGGA GCCAGGCAGG CTAAAGAGCT GTGGTTGGAG
5051 AGAGGAGAGC CTGGATTAGG GAGATTCCAC AAGGAAAGGA TCACAGAGGA
5101 CAGCAGCAAA GGGCAGAGCC CAGAGCTGTA TGGAGGGAGG ACGAGGGTGG
5151 GCCTACCAGG ACACGGCAGC TCCAGGCTCC TTTTAAGGAG GAATCCGTAA
5201 GTGGTTGTTA AGCTTGACTT CAGGCCTGGG GTGGGGGCAG GTTCTCATTG
5251 TCTTCAGCTC CTGTTCTAG GCCCGGTCTT ATGGCTTTTT AACCAAATAA
5301 GGCCAAGGCC AGAAAACCC CAGCAGCAAT AAAAGCAGAA GGCTGACCC
5351 AATCTGGGAG GCTGGGTTTC CCTCCTAGGT CGGCCACACC ACCCTCTCCC
5401 ACCCTCCCTG CTGGGGAAATG GACCTGCAGC TCCCCCATGT GTCTGCTGGG
5451 AATCCTGAGA GAGTGGGCAC CCCTGTTCAC ATGCCTGCTC CCTGTCTGCT
5501 GCCTGCCCTA CCCCCAGTCCTT GGGCTCAGGC TCAGTCTTGT GTGCCATCAG
5551 CCCCCATCAGG AGAGCAAGAA TGGCAGGAAG AAGGGATGGG AAGTGAAGAC
5601 AGTCGTAGCA GAGGGCTCAG TTGCTGGGTG TTGTGCTTGG AGCTAAGGAG
5651 ATTGTCAAGAT TCTGCAACAG CTAGTGCAC ACAGATGCCT CTAGTCCAGG
5701 TGGTCAGGTG CTGCCAACAG GCCTGGAGCA AAACCTTACA GGGCCCTACT
5751 GTGCCAGGTG TAAACTCTTT AACTGCTTTC CTAAGGATGC CTGGGGGGTT
5801 CTAGGGGAGC AGCCAGGGAC CGTGGGATAGT GGGGGCATTT GGGGACTCAG
5851 AAATAGCCAT ATTGTAGATA TTTCAATATT TTACCAACCC TATAGCCATA
5901 CTGAATATCA GCCATGGAGG GCCCTTTCCA AACTGTCCAC TCCCCCTCCA
5951 TTACATAACA AAAGCAGCCA TCATTTGCTC TTTCTTCAA CAAACGTGTA
6001 TTGAGTACTG AGTTGGAGCC TAAGCACTGG GTCAGGGAGA GCCCTGTCAC
6051 CCTGGGCTTC GAGGCAACCA CTTCCAGGCT TTACCCAGA TCAGGCAGAG
6101 ACCCCCCAAA GGAGGCTGCT CCACCCAGCA GCATCTTAAG CTGAGTGGGC
6151 TCAGTGCTC CTTCTAGAC AGAGCCCAAT GGAGCAGTG CACTGATTTG
6201 CAGAGGTGAG CAGATCCAGC CTCGTGGAAC CAGTAGAAGC CCAGCCCTGG
6251 TGAAGCTGTT GCTAAGCAAC ATTGGAGCCC ATTCTGAAAG GGTCCATCTG

FIGURE 3, page 2 of 33

6301 TTGGCCAGCC CAACTTCACT GTGTTCTGAG CATTCTGCAT TCCTCAGTCC
6351 CATCTGCTCC CTCCCATGTG CTTGGAGTG ATATAAAAGT CCACCAGCAT
6401 CTCAGTGTGA GCTGACAGGG GCCAGGCAGC ACCTATTTT GTCCTAGATG
6451 TGTCTAACAA TAGAGGCAAC AGGCAACAGG CAAGACGCAG TGGGGGGCGG
6501 GAGGCAGGAG GCCGAGATGG CTGTGAGCAT GAGCTTCTC AGCCCTCTCC
6551 CTTCTCCCAT CGCGAGTCTA ACTGCTCATA CGTTCTGTGT GCCAGGTAGG
6601 GTGACTTAAC AGCACGCCAT GGATTTCTGT TGTAGTTCA AGTGGACAA
6651 ATTCTTTTAC AGACAACTT TGACTAGCCT TCTGTGGACT GAGCCTATAC
6701 TCTGCCTAA TGGGCTCTCT GCCCACTCCCT TTCCCTAACCC CAGGGCAGCT
6751 GGCTGAACAC CTGGTCCTT TCTTAGGTTT CATTCTTTT GACCTCTCTG
6801 AAGCCCTGT CAAAAGTCAC CACCTCCCCC TTGAAATTCA CTCCCTCCTG
6851 GGTTTGTGGA CACTAAATCG CCTTGATTT TCTGGTCTTC TGTTTGCTTG
6901 CCTTTAATGA CCCTCCTCCT CCCTTCCCCC AGTCTTGAAA ATGAGATAT
6951 TCTCCAATT TCATGTCCTCC ATTCTATTTT CTTTCTTTT TCACTCACTT
7001 TTTGAAACAG GGTCTTGCTC CGTCTCCCAG GCTGGAAGTG CAGTGGCGCA
7051 ATCACAGCTC TCTGCAGCTT TCAACTCCTA GGCTCAAGCC ATCCTCCAC
7101 CTCAGCTCC TGAGTAGTTG GGACTGCAGG CATGCACAC CATGCCAGC
7151 TAATTTGGT TTATTTGTT TGGTAGAGGG GGGGTCTTGC CATTGGCT
7201 CAGGCTGATT TTGAACCTTG GNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
7251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
7951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
8951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN
9401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN

FIGURE 3, page 3 of 33

9451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 9501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 9551 NNNNNNNNNN NNNNNNNNNN NNNNCAATAA ACCCAGCTAA AAACAAGCCC
 9601 AATAAAACCC AATAAAACCC ATTAGACAGG AACATAGGAG TTGGAAAAAA
 9651 AAGAAAAGAA GGGGAGGGGG AAGAAAGCCC TGAGGCACCC CGGCTGCCTG
 9701 TCTGCCACAA CCCTGGGCTG TAATTGTTCT TGCCATGGCC TCAGTCTGCA
 9751 ACACATTCTA GTGTCCTCTT GACCTCTAGC CCTCTAGCTC TGCTCCCTT
 9801 TCCCCAACCT GTAGATCTTG TGATCAAATA GATTCAATGA AACACATTGT
 9851 CCAGTTGCAC TCGCAGCACT TCCAAAAGG TCAAGTTGT CCTCCCTCA
 9901 GTGCCTCCA TTCTGGTCAC GGTAGGACTG ACTCCAGCCC CTGGACCCTA
 9951 AGCTGAGTCT GGGCTCCTT GACGTGCAGG GAGAATGCCA CTGAGTCTTG
 10001 TCTCTGAGGA CCCTACCTCT CCAAATCTTG CCTCAGTTCC TCAGCAGGTA
 10051 CTACACTGAC TGGCCATGCC ATTCTCTGAT GCTTCACTGC CTCAGCTTCT
 10101 CAAGTCTGTC TCCCCACCTG AGCCAATTGT GAGTTCTCT CTCTCCTCCT
 10151 CTCATCCTGG CACCTAGAAA TGCTCTTA CGCTTGAGCT GCTCAACCAG
 10201 CATGGGTCACT TTGTTTATAG CATGTCCTCA GATGCCCTC TTTGTTGGTG
 10251 AATGCTCAGG GAATGCTTAC TGTAAACCC AGACAAGCCC AAGTAGCTAC
 10301 ATGGACCTGC CACCATAAGC CCTCTCTGT CTTATGCTGT TGTAGAGGGT
 10351 CCAGGGCTCA CTTCTCCCAC TTGGCCCTGA GTACCTCTCC TTGAAAGGAT
 10401 GTCAGGGCT GGGCGCAGTG GCTCACGTCT GTAACCCAG CACTTTGGGA
 10451 GGCTGAGGCG GGCAGGATCAC CAGGTCAGGA GATCGAGACC ATCCTGGCTA
 10501 ACATGGTGA CCCCCCGTCT CTACTAAAAA TACAAAAAAAT AAAAATAGCC
 10551 ATTTGTGGT GCAGGTGCCT GTAGTCCCAG CTACTCGGGA GGCTGAGGCC
 10601 GGAAAATGGC ATGAACCCAG AAGGCAGAGC TTGCAGTGAG CCGAGATGCC
 10651 GCCACTGCAC TCCAGCCTGG GCAACAGAGC AAGACTCCGT CTCAAAAAAG
 10701 CAAGCAAGAA AGAAAGGATA TCGGTTACCT GTTTCAGACA GGAATGCTGA
 10751 GACCAGGGAA AGGGGAGACT TGTCGGGTGIC CTCAGGGAAC CAGTATCTGA
 10801 GCTGGGGCT GAGAGCTCTG TGTGGGTGGA CTCTGTCCTC CCAGTCGCTG
 10851 CTGAGTCCT CTCTCCCTT CCGCTGTCT GACCAACAGG GTTTTTGTTG
 10901 GCCTGACCTC CAGTGTGAGG AACGGAACCA GGCAAGAGGC TTGTCCAGTC
 10951 AGCTCTGGCC CCAGTTTGGC ATTCACTCACT GTTTCCTAA CCTGGAACCC
 11001 GTCCCTTCT CACTCTGGGG CTCAGCAGCT GCTCATCTAT AAAGTGGGGC
 11051 ATTTGGGGT TGCAAGTC GTCATCTACA ATGCCAGGCA AGAACATGGT
 11101 TCGTACATG TGGTCAGGTA TGAGACAGGA TCACTTTCC AACATTCTGG
 11151 TTTTCCCTT TTTTTTTT TTGAAACAG AGTCTGCTC TTGTTACCCAG
 11201 ACTGGAGTGA ACTGGCATGA TCTCATTTA CTGCAACCCCC TGCCCTCCAG
 11251 GTTCAACAA TTCTCATGCC TCAACCTCTC AAGTAGCTGA GATTACAGAT
 11301 GTGTACCA CCGGATTTT TTTTTTTTT TTTTTTGTAT TTTTAGTAGA
 11351 GACAGGGTTT CACCATGTTG TCCAGGCTGG TCTTGAACCTC TTGGCCTCAT
 11401 GTGATGAACC CCCCCCTGGC CTCCAAGT GCTGGGTAA CAGGTGTGAG
 11451 CCACTGTGCC CAGACTAATT TTTTTTTT TATTGAAACA GAGTCTCACT
 11501 CTGTTGCCCA GGCTGGACTG CACTGCCA ATTATAGCTC ACTGTAACCA
 11551 CAAACTCCTG GGTTCAAGCA ATCTCTGC TTCAGCCTCT TGAGTAGCTA
 11601 GGATTACAGG CACATGCCAC CATGTTGAGC TAATTATTT TTTTAATTG
 11651 TCGTAGAGAT GGGGCTTGC TATGCTTCCC AAGCTGGTAT TAAACTCTTG
 11701 GCCCAAGCA ATCCTCCTAC CCTGGCCTCC CAAACCGCTG GGAATACAGG
 11751 CATGGGCCAC TGTGCCAGCC TGGTTTTTC TTCTTGTTC CATTATATTG
 11801 TCACATTTC AGACCATGGG CTTACTACTC CACTGAGCAC ATTTTGTGAG
 11851 AGTGCTACA GCCCTGGGC CGGTTGCTGT TTCTGATCT CAGTCTTATC
 11901 AACTTGATCT TGCTTTGCTG TCATTTATAC ATTTCTCAT TAGCTTTCTC
 11951 CCCATTCTT CTTTGTCTGC TTCTTCTTC CTTCTTAAC TAACTCCTCA
 12001 CCTGCAACTG GGGGGACTTG GATTCITGAC TGGGCTTGTG TGGAAACTGA
 12051 TTGAAAACA GATAGGTAAG TAGGGAATGA GGAGGGTGTG TTACAAGAAA
 12101 AAAAATGA CTAAGATACA GGAACCCAAAC CTAAAGAGGA AAAGACATAC
 12151 AGTTCAAAGG AGGCAGGAAAG AAAACATTA CAGATACTCA AATATATTGA
 12201 TAATCATAAC ACTTTCTGGA AGATAAAAAA AATGCTGAAA CATGAATCCC
 12251 TTGCTAGAGA AATTACAAAG CCAAGAAAAAT AGATAGGTCT GAGGATTAGG
 12301 GAGCTGTTCA GTTGCTAGGA GGAACACAAA AGCACAGACC CCAGACTACA
 12351 ATGGGTATGA AACCCCTGC ACAGCCTTTG TTGTCCATCC CTGCAAAG
 12401 CTGTTATGTA AAAACCTCCG GGGGAATGAA TGAAATTATG TTTATACAGT
 12451 TCTTTCTATA TAAGTGCAGA AGAATCATGT TAAATAAATC TACAGGGCAG
 12501 GATTGTTAGT TTTTCTCCTT CTCAAGAAA CTTCAGTGCT GTCAAGATAAC
 12551 TTCTCCATGT GTTTTTTTT TCTCTTAGAA TGAGCCCTTC CTGAGAACCC

FIGURE 3, page 4 of 33

12601 CTCGGAATAG TAACTACACG TACCCCCATCA AGCCAGCCAT TGAGGTAAACC
 12651 CCTGACTCAC ATCTGCCTCT CTCAGACACA AACCATTTCC ACCTGCCAGG
 12701 GGCTCGGGTG TGGTACAGGT TTCAGAGTAT TCACTGAAGC AGAAATGTAC
 12751 TTCTTACATA CTGGGGATTG GAATGTACAG AAAAGGCTCC CGGACCACGA
 12801 AGCCCCAGGA TTGTCCTAAC ATGTTCTCAA GTTGCTTAC TGACGTCA
 12851 CCCCAAGCAG AGGAAGTGT TATGGATCGA TTTTCTTGA CCTTGGCAAT
 12901 CCTGGGCTCA CAGACGTGGT TACTGCTTAG GCAGCTCAGC CTCTCAAGAG
 12951 GGAGAGGCAG CTGGTGTGAT GTGGCGTTGA CTTCTGGAA GGTGGAGGCT
 13001 GAGTGGGAGG GAACTACAAT TCTGGGGATG GGACCCAAA GGAAGTGGAG
 13051 GCACGTTGTT CATGTTCTG TGGGCCCTA GGCCTTGTGTT GGTTCAAGTC
 13101 AATCATTCTA GTGCTGAGGA TTCAGAGCC ATGGTTAATT CCATTGGATT
 13151 AACCATGTCT GTGAGCCTAG GACGGCCACT GCAAAGACGG CCTGGAGGAC
 13201 CCCGGACTAT ACCATGACTG GCAGTCAGGC CTGGTCCGGA TCAGGTCTGT
 13251 TGGTCACCAG GATGGGGTTT GACCCGCAGT TTCAGTTCA CACCTATATT
 13301 ATATCCAGTC TCATGTTAGG GGCTAGAAGG CATGCAGAGA AGTATCGAAC
 13351 ATGGTCCCGA CCAAGGGAAG TGAGAGCCCA GTAGAATTTC ACAATTATTG
 13401 AGCACATACT ATGTGCCAGA CACTATTCCA GGAAGACAGA AATGTTAAC
 13451 AGACAGATGG ATCCCAGGCC TCACGTAGCT TACAATCTAC TGAGAAAGGT
 13501 GTCTTATATA CATGGCTAGG CATGGTCATT TCAGATAGT ATGACAGCTC
 13551 TGAGGAGCGT GATGGGGCTG GGGCAAGGGA GGCIAATTCA GGTGCACCAT
 13601 GCAGGCCAGG CCTCCTGAG GTGAGATTAA AACTGAGACA TGCATAATGA
 13651 GGAGACACTT GCTATACAGG GAGCCAGGAA CACAGTCCC GGCAGAAGGA
 13701 CCATGGACCA CACAGGCTCA GAAGTGGGAC TGTGTTGGGT GTATTGGGG
 13751 AAGAGAAAGA AGGTTCAGAGT GGCTGGGGGC ATGAGAATGA GGTGGAGAGT
 13801 GGGGGAAATG AGATCAGGAG TGCCAAGGAG CCAGATCACA CAAAGCCTGA
 13851 ATTACTGAGT AAAACCCTG GATTTCAAGT GGAGAAAGAT GGGAAAGGCAT
 13901 TGGCGGTCTC AGGAGAGAGT GACATGATCT GGTCACGTC TTTCAAAGAT
 13951 CTCCCTGACT GCTATGTGTA GAATGGGTTG GCCATCAGCA GGAGTGTATTG
 14001 GGGAAAGACA TTTTATAAGC CAGCTGAAGA AACTAACCCA TATGAAATCA
 14051 TTAAGAACTA TTGGATGCTA AGCTCTGGGG TGCAAGCAAT ACCAGATTGC
 14101 TGGCTCGGGG TTATGCTGTG TCCAGCCTCT CTGAATTTC TCAGGCTCAC
 14151 GTTAGGCCAG TGGAGGCTTG TCCTCATTGA ACCAGTGACC AAATTCCCTG
 14201 AGAATTGAAA CGTCAGCTGC ATCTTGTGAA TCAGGCATTCTTCATTAT
 14251 TCATTACCT ATTGGATGCC TATGTAGAGT GGGCACTGCA CTAAGTGCTC
 14301 GGTAGACAGT GGTGAGCCGA ATGGGTCTGG ATCTGCCCTC TTGGTTCTC
 14351 AGTCTCATGC ATCTTGCTT TTGCTGCTGG AAGAGCTAAA AAATCCAGAG
 14401 CTAGAAGGGC GTGTGTTGT TTAAACAGCT TTCTACTCAA AGTAACCACA
 14451 GAAACAAAAT TCTGTCATCT GAGGTAAACGT GAATGAGCCT AGAGGACATT
 14501 ACGTTAAAGT AAATAAGTCA GGCACAGAAAA GACAAATACT ACATGTTCTC
 14551 ACCATATGCC GAAGCTTAAG AAGTTGACTT CACAGAACTA GAGTATAAAT
 14601 AGTGGTTATT AGAGGCTGGG AAGGGTGGAT GGTGGTTGG GAGTAGAGAT
 14651 AGCAGAAATT GATTAACAGA AAATTACAGC TATATAGGAA GGAGAATTTC
 14701 TAGTGTTTA TAGCACAGTA GGGTGAATAGTAAAGT TTACCATATA
 14751 TTTTCAAATA GCTAGAACAG CAGATTGTA ATGTTCCAA CACAAAGAAA
 14801 TGGTAAATAT TTGAAGTGA GGATAGGCTA ATTACCCCTGA TTTGATCACT
 14851 GCACATTGTA GAGATGTATC AAATATCAC ACTATGCCTC ATAAGTATGT
 14901 ACACCTAATA TGTCAATTAA AAATATAAA AGCAAAACTA ATAAAGTGGC
 14951 CACAAAGAGG CTTTACCTGG GAGCTTTTA GAAATGCAGA GTCCCTGGGA
 15001 CCACCCAAA CCTGCTGAAT CAGAACTGTC AGCTTAAGAT CTTCAGGGGA
 15051 TTTGGATGCA CTGATTTGG GTGTGGTGCAT TGGTTCTTCC CTGTCAGCG
 15101 ATGAGCACGT TTCAATTCCA ACCGGATCT GTTAATCTAC ATGGAATATG
 15151 CTTATCTCTG GTTCAACAC TATCTGAGAT ATATCTCATG TGCTGATGGC
 15201 TGAATAACTT TTTACGTTGC ATTTTCTGTG AGTATTGTC ATCTGCACAC
 15251 AACGATGCTC TTGAGTTCAT TAAACCTTTA AACAGAAGAA ATCCATCAGA
 15301 ATGATGAATT GAGCAATCCC TTGGGAAAAA ACCAAATTCC ATAGGATTAA
 15351 GCAAATAATA TTTAAAGAA GTTCCATTTC TGCTCTCTCA TGATAGGAAT
 15401 ATTTCAACAA GTCTTATCTT CATCATCTGA CTGAACAGAT GAGATGAGTT
 15451 TTCACTAGCAT CTGGCAGTC GACTCCTGGA CAGTCATCT GCTGGTCAAG
 15501 CCCTACTCCA TACTCAGTAT GCATATAATT GAGACTTTGG GAAGATACTC
 15551 AATTTCCCCC CAGATTCTG GTACTAATCA TTTCTATGCC CTCTGCTTCC
 15601 CATCCCACTC CTTCCCCAG CACCTGGAAA ATATGTTCTG TATTAGAGAC
 15651 AAAGAAAATT GACTAAAAGC ATCCAGGGTT GCTTACATCA ATTAAAAAAC
 15701 ATATAAGGAA TAAGGCTGTT AAGTTAAATA TGCAAAAGA CATACAGGTA

FIGURE 3, page 5 of 33

15751 TCCAGAAAAG ACAGGCAGAA ACCAGGAGCT TTACAATTTT AAAATATTT
 15801 GTGTTATAT TCTAAAATA TTTAATTAT TGTCTAGGTT CTACCATTAT
 15851 AATTAGTGTCA AGTTAGCTTA ATTTTATAAA ACACACATAC CTGTAATCTC
 15901 ATGTTAGGCA TCCAATGCT GTGTTCCCTT GGGAGACCCA CCTGTGTTAGG
 15951 ACTTCATGGT TTTCTCCCT GCTTTGGGGC AGCCACTGGC TCCATTCAA
 16001 GCATAGATAT ATGGGATAA GAAAGGTTGT GTGTTGGTGC ACATGTGGAG
 16051 ACATGCACTA TGGGTTGTGC ATAGGGTAG CTAGACACAC CCATTTCTCC
 16101 CCCCTTAATT TCCCTCCTAG CCCACCTATA ACTCACAGTT CTTCCCTCA
 16151 CATGATCCTG TATGGTACT CATTCTAGC CTCCATCAA AATCCCTAG
 16201 CTGGTTCTC TTGGGCTGAA GCTTATCTCC CTGCACAATG AGTGTGGGC
 16251 ACTGAATCTT TTCTCCTGTT GATTAGAAC TGGGGCAGTG ACTTCCTGTG
 16301 TACAGAGTGG AAGGCTTCCA ATAGTGTCC AACCTCTGGT GAGTGAAAAC
 16351 ATCATCATCT CCTCAATTA AGGGCCTTGC CGAATATCAG GTTGTGGGA
 16401 GACCCTGCAA ACATACCTG GAGCTTAAG CAGGACTTGC TAATTCCCCT
 16451 GCAGTGCAGA CCTAGATCCT GCGGCCCTGCC GCCACAGCTG GGCTTCCATG
 16501 TGGAGGTGCA CAGAGCTCTC CATTGGATGC TACTTCTTGT CTCCTTATAG
 16551 TCCCAGTGGC AGTCCCTTAG GCCTCCCTGC CCAGTGAGGC AGGTAGAGTC
 16601 AGGGATTGGG ATCTACCTGC CTGTGCTACA TGACCCGTCA GCTGGAACATT
 16651 TCCTGGACCA CCCCAATGTC AATCAGGCTC TTCTGAGGGT GGATGATAGC
 16701 CATGAAACCC ATTCCCTGCA GTGCCCTGGT TGGTCTGAAT GAATGGGAGG
 16751 GGCAAAACTG CTAAAGCCTT AAGCTGAAAAA TAAGTACAAT GGGGAGCAGT
 16801 GGGACAGAGT TATAGACTTC TGGTAAAATG TGTACTTTAA GAGGTAGATA
 16851 CCCCCAGCCC CCACAAACAC CTCTCTGCTT GTCTCCCTA GTCCACCAGC
 16901 TCCGACCAGC AGACATCAA GTGGTGGCCG CCCTGGGTGA CTCTCTGACT
 16951 GTGAGTAGTG AGCCATGAAC CAGGATGGGC AGCTCAGAGT CCAGCCAGGC
 17001 CCTGCGCAGA ATCTGTGCTT CCCCAGCATT GGCTCCGCTT TCAGTGTGA
 17051 GCCCCGTGTTA CTGAGGGCCT ACCCATGTCA GGCACTGAAA CACAGCCAGG
 17101 AGATGTAGAA TGCCCTGTCT CGCCACCTTC CCAGTTCTGC TCAAAGCCCC
 17151 CTCGTCATG AGGCCTCCCC TCAATTCCCC AGGGAGAAC AATCCCCGCC
 17201 TTCCCCACTG TTCACAGGCG TTTTGTGGT GTGTGATGGC ACTCATTCAA
 17251 GTCTGCCCTGC CTTCATCAGG GGACTCATCT CCATCTACCC AGACTCAGAG
 17301 TGGCAGGTCT TACACACACA CTGCCCCATG CTCCCTACTC CATTIAAGGA
 17351 CATGTGCTT GGGCAGAGG GAGCCCCGGT CCTCACACAT AGCACAGTCT
 17401 TGCTAAGTGA ATTGTGTTCG CCAATTACT AGCCATTGTT GTGTACACCA
 17451 ACACTCTATT AGCAATTCTA AGGGAAATGA GGTATGAAAC ACAGTCATAG
 17501 CCCCCCAGCA ACCTGTCCTGG CTGGAAAAAC AAGAAACGTA CACAGAAAGA
 17551 AATGCATAGT CATAGATG ACATATAGGA CTTGGATGTT TTATTTTTAT
 17601 TTTTAACCT CTAAGTTCAG GGGTACATGT GCAGGTTGTT TACACAGGTA
 17651 AACTGTGGC ATGGNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 17701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 17751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 17801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 17851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 17901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 17951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 18801 NNNNNNNNNN NNNNNNNNNCT CCAAGTTAA TATGGCCTGG AGAAAAGCCG
 18851 GTACTATTAGAAAAGGC AAATCCAGGT CCTAGCTGCT ACCCCAGGGC

FIGURE 3, page 6 of 33

18901 CAGAGGAAGG CTCTCCAGT GCCCTCAGCC TATACCCAG CCCTGAACCT
 18951 TCTTTTGCT TTTTACAGAC AGCAGTGGGA GCTCGACCAA ACAACTCCAG
 19001 TGACCTACCC ACATCTGGA GGGGACTCTC TTGGAGGTGA GGATGTTCTT
 19051 GATGCATGCT CTATTGATGA TGCTCTCTCA GAGAGGTGTG AGTAGTGTGT
 19101 TTCCCTGTAC CCCTCCAGGG ATGCAGTTGG GTCCCCAGGT CCCAGCGCTG
 19151 AGACAGGAGA CTCATGCTT GCATTACCCC TGAGGGTGAT GGGAGAGACG
 19201 CCCCAGGGGC CCAGAACCCG GTTCCGGTTC TGGCTTGTC ATATGTTGAC
 19251 ACAGGGAGCA GCATGTTGGT GTGAGTTAA CAAATATGCT TTCTCCTCCC
 19301 CAGCATTGGA GGGGATGGGA ACTTGGAGAC TCACACCACA CTGCCAGTA
 19351 AGTAGCAGCC CAGAGAGGCA CCATCACTGT GGCGTCCCTC CCTGGGGCCA
 19401 GGGCCTTCCT GCTGGAGGAG GGGAAAGAGGA GGTTATCTGC AAGAAGGGAA
 19451 GTCAGCCAGC CCTGAAAAGC CCCAGACTTC CTGTGTCCTC CCCATGTCCC
 19501 CACCCCTGCAT GCTCATCTCA GTTACTGTGA GGGTCCTGCA GGCTCTCAC
 19551 TGTGCTCTTC TCCTCCTCCT CTCCTCTAA AGACATTCTG AAGAAGTTCA
 19601 ACCCTTACCT CCTTGGCTTC TCTACCAGCA CCTGGGAGGG GACAGCAGGA
 19651 CTAAATGTGG CAGCGGAAGG GGCCAGAGCT AGGTGAGTAG ATGCCGTACA
 19701 GGAGGGCGAG NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 19851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 19901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 19951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 20951 NNNNNNNNNN NNAAAATCA AAAATTAGCT GGGTGTGATG GTACACGCCT
 21001 GTAATCCAG CTACTCGGGA GGCTGAGGCA GGAGAATCAC TTGAACCCAG
 21051 GAAGTGGAGG CTGAGTGAG CCAAGATCGT GCCATTGCAC TCCAGCCCGG
 21101 ATAACAAGAA TGAATCTCCA TCTCCAAAAA TAATAATAAT TAAAAATAAA
 21151 TAAAAGATAC AAAGGAATCA AAAGATGAAC TCCCTGGCCA CGAAGAGCTT
 21201 GCACTCTAGG TAAGGAGGCT AAACAAATGG GAAATAACTT TCTGAAAAG
 21251 ACAATGCTGG GTATGGCAAC AATGCAGTGC TTGCACTGGA GTACAATTAA
 21301 GAGAACAGAA GAGCACACAG TATGAACTGC ACTGTCTAAA GACAGATGCA
 21351 GACCCAGAAG GGACCCCTGA AATCATCCAG TCCAACCTCT TCCTTTAAAA
 21401 GATGGGGAAA GTCAATCCTA GCAAGATTCA GCAACTTGTA CAAGCTCAAC
 21451 AGCAACTTGG TAGCAGAGCT GAAAGTAGAA CCACTGGTCC CTGGGGTAAA
 21501 AAAGGAAATG CAAGATGTGT GGATCAGGGGA GCCAGAGAG GAGGCTCAAG
 21551 GGAAACTAGG ACTTGGTCTG GGCTGAAGG ATGGGAAGAA GATGGCTAGG
 21601 AAGAGGGGAA GAAGCGGCAT TTGTAACCTC CCCTCTTAC CACGAGGGCT
 21651 TATTGCCAT GGATTCTCTT AGTCACACT TGAACTGTG AAAAGGTAA
 21701 AGGCACCTCT GTGGTCACCT TTGACCAGAA AAGTCTTCTT TTATAGCTTT
 21751 CTGGTATACT CATCAATAGC AATAATGTAT GGGATACAAT CCTAGATCTG
 21801 TAAATTCTCC TTAATGAGAA ACAAGGGTAG GGATGGTACC ATGTGTGTTT
 21851 GGCCACGTAC CTAGCTTACG GTGGCCACCC AAAGATTTTC AGTGGCCAGC
 21901 TCGCACTGGT TGCTGCTTTT ATGGCTTCTT CCATGGACGC TTTCATTGGC
 21951 TATATCCCCT TTGCTGACCT AACTTCTCAC AGACATTCTT TTAAACACAG
 22001 CTTTATTGAG GTATAATGA CATGCAATAA ACGTCACATG TTAAAGGAT

FIGURE 3, page 7 of 33

22051 ATAATTGGC TGGGTGCGGT GGCTCACGCC TCTAATCCC GCACTTGGG
 22101 AGGCCAAGGT AGGCAGATTG CCCAGGAGTT GGAGACCACT CTGAGCACT
 22151 TGGTGAACC CTGTCCTAC CAAAAATACA AAAAATTAGC CGGGCATGGT
 22201 GGCATGGACC TGTAGTCCC GCTACTCGGA AGGCTGAGAT GGGAGGATCA
 22251 CTTGAGCTA GAGGGGTTGA GGCTGCAGTG AGCGTGTGATC ACACCACTAC
 22301 ACTCCAGCCT GGGCAACAGA GCAAGACCC GTCTAAAAA GGATACAATT
 22351 TAACATTGTA CCTGTGAAAT CATCACCACA ATCAAGATGA AAAATGTGTT
 22401 TATCACCCAC AGGAGTTTTC TCAGGCCCC TGGTAATCTC TCCCTCCTGC
 22451 TCCTTCCTGT CCCTACCTCA CACCCCCAGGC AACCACTAAC CTTCTTCCA
 22501 TCACAATAGA TTAGTTGCA TTTTAAAAA TTTTATATAA ATGGGATCAA
 22551 AGAGTATATA CTTTTTATCT GACTTATTGA GCAAAATGAT TTTGCGATGC
 22601 ATCCATGTTA TTCGGTATAC CAATAGTTCG TCCCTTTA TGGCTGAGTG
 22651 TAGTGTTCGG TTGGCATTCA TATCGCTCAT CCAGAACACC AAATGGTATT
 22701 GTTTTATTTA TGCGAGACAT CAGGGGATGA AGGGAGAACT AATCCTGTCC
 22751 ATCCTGGTT ATTGGAGAGG GAGAAAAAAA AAAGTGAGGA GATGGGAAT
 22801 GGTGCGGAAA TCTAAGTAAC CACAGAAAAG AAAAACAAAA GGATTAAGG
 22851 AGCAGAGAGC AGGGCTTAGA AGTAAAGGTT AAAGGAGTCA TTAAGCCTGG
 22901 AAAGGAGAAA ACTGAGGGAT AATTGTGAGC TGTGACTTT CTCAAATATA
 22951 CAAAAGGTTA TTTTAAAAC AGGCAACTGA AGAAGAAATG AACAGGCTTG
 23001 GCTTACGAAG AAAGAGCTG AGGAAGTATA AGGGAAAGTC CCTGAGGGGA
 23051 GGCTTGACGG GATCCCAACC CGAGTGGCCG ATCAGACTAT TGGGTGGCAG
 23101 GGGCTAGATC AATGTGGCTC CAGGGTCCAG GGCAGCCATG TGATTGTTAC
 23151 TAAGCTGAGA TTTCTTGAGA ATGGAATGAC CTTTGTACTG GTAACATCAT
 23201 TCTTCTTGAA ACACCTCTC TCCTAGGCCA AAATCCCATG TCGTGAGTCC
 23251 TCGCTCCTGA GCCGGCACTA ACGCCCCCTCT CTCTACCCCC CACCTAGGG
 23301 CATGCCAGCC CAGGCCCTGGG ACCTGGTAGA GCGAATGAAA AACAGCCCCG
 23351 TGAGTACAGG CCCCCCAGGCC ACCCCTGAAA GGTGCCCATC TCCTGCTGG
 23401 TGGGGAGGGG ACAGCCCCAT AAGGGTCCCT CTCACCCACAG CACTTCCTGC
 23451 TTTGGGCTAG CCAAAAGATC CTCGGAGAAAG CAGTCCTTAC CAAGGAGGCG
 23501 CCTGCCCTGG CCACACTCT AGACGCAGGC TGTGGCACCC CTCACCCCCAG
 23551 GGCCGGCTGC GGGAGGGAA GGTGGAACAG GGAGTTGGCT GAGGTGGTGG
 23601 CCTTGGCCTC TGACAGCTC CTGCTTTAAC CAAGAGGTGG CTTCCCAGAG
 23651 CCCTATTATG TAAATGCAAG GTTCTAAAAAA TAGGCTTCTC ATTCCAATCC
 23701 AGTTCTGGCT CCTTCCCCCTC ACCCTGCCCC TCTGAAACTT CTCACTAGCA
 23751 CTTTTTTTT TAACCGTTCA GTGTTTATGC CTAGGAATTG AGCTCCCGGT
 23801 GGGATTCTTA TTATGGAGGT GGCAAGTGG AAAGCCAATC GCTTAGAGGG
 23851 CCTCCCAGCC CCAACCCCGC TTCTCAGTCC ACGCTGGGCT CTTCCCTCAG
 23901 TCTCCTTCCC CCGACCCCTAA GAACTCATCC CAGGGGCAGC TTAGGGCCTT
 23951 TGCTTCTAGC TGACATCCTT GCCTACAGCT CCCTGGAAGG CCTTCATTTG
 24001 GGGGGACGTG GTAATCCCCT CGGCATTAA TGGGCCAAGG ATATGTGGG
 24051 CACATCCACA TTCTACTTCT CCAGGGACAC AACTTTCTTA AGATTTCAAG
 24101 GGGAAAATAG CCCTCCCTTG TGTAAGCAGA ACCCCGTCCC CGCCAGCGC
 24151 CCACCGCCAA AAAAAXXXXX CATCCTCTC GTGGAGCACC TTATCCTAGC
 24201 ACCAATTGAG GGCTGGGAAG CCCCACCTTG TTGTTTTCT TTTTTTTTTT
 24251 TTTTTAGACG GAGTCTCGCT CTGTCACCCA GGCTGGAGTG CAGTGGTGCG
 24301 ATCTCAGCTC ACTGCAGCAT CTGCCTCTG GGTTCAAGTG ATTCTCCTGC
 24351 CTCAGCCTCC CAAGTAGCTG GGATTACAGG CACCTGCCAC CAGGCCAGGC
 24401 TAATTTTGT ATTTTAGTG GAGACGGGGT TTCATCATGT TGGCCAGGCT
 24451 GGTCTCGAAC TCCTGACCTC AGGTGATCCA CCTGCCCTAG CCTCCCAAAG
 24501 TGCTGGGATT ACAGGTGTGA GCCACCGTGC CTGGCCAAT TTGCTGCTTT
 24551 TCTCTGTTAC AGTATAATAA AGACAAAAGG CACTTGGAGG GCGGGGGGGC
 24601 TGGCCCACTA CCAACATTG TATGTGCCCTC CCACCAAGGC CTAACACTCAG
 24651 GATCTTCTGT CCCCTCAGGA CATCAACCTG GAGAAAGACT GGAAGCTGGT
 24701 CACACTCTTC ATTGGGGTCA ACGACTTGTG TCATTACTGT GAGAATCCGG
 24751 TAGGCCCTCG ACCAACCCCCA TGGGGACCTG AGAAGGAAGG TGCTGACCTC
 24801 TGGCAACACC CTTGCCCATC CATCCCTGGC CCTGCCCGA GCTCCTCGCT
 24851 CATGGGAACC ACATTGCCC GCTGCCCGAG GCCCTCCCTG GTTACACAT
 24901 GCCAGGAAG GCCCAGCCCT TTCTACTGCC TGAGCGACCC CTGGAAGAGC
 24951 AGGTGCATTG GTTCCCAAT TCCAGAAGTA AGGCAAGGT GGACCCACTG
 25001 TAGGCACTGC TGAGGTGAGG CCTCTCTTAT CCACACAAAT ATGACCTCTG
 25051 GTACCAGATA GGGGACTAGC CATCCTCACC CCATCCCTGC CCTGTTCAT
 25101 TTTGAGGAAG GGCAAAACAA TGTTCTAAAT GGGGTGGAT GGGTCATCAC
 25151 GAGTTAACCA AACCTCAGTG GTGCCCTGG GAGCCAAAC CTGTTCCCTGA

FIGURE 3, page 8 of 33

25201 TGTTTCCATG GGGAAATATG TCCAGACTCA CAAACTCTG GAAGTTGTAA
 25251 CTTCCAAAAG TTTTTATTTC GGAATACATC TTGTTCTAAG TTGGGGACAT
 25301 TCTGTATTT ATTGGGCCGG GAAAAGGCAG CTGCCAGCC TCAGAGATTG
 25351 TGGTCGAATG TTTGACAGTC ACAGATGGTG GTGTGGGGAA AAAGGCATTT
 25401 ATCACTCCTT GTGGCTCACA GAGGCCAACT ACCAGAGGCT TGATGAGAAG
 25451 TATCCTCCA CACACAGGAG TTGGTTTTC AAACCTCTTC CCCTCAGTTC
 25501 TCTCCCTGAC CACCCACACT CTAGAGGCAG AAGTGCACCT AACTCATGGG
 25551 GATTTAAAGTA TTGCTGTGTT CTGGACTCCG GGGATACCTG GACCCCCAGAG
 25601 CTGCGTGGAC ACTGTGGACG CTGGCGTAGG GAAATGCCCT CTACTAGTCC
 25651 CAGGATGTGT GTTCTTCATG GGAACAGTC GGTTTATACT TCCGAGAGCG
 25701 TAGTTTAGTT GAAAGGGCTG GGCTGCCCG ACTAGGATTA ACTCAGACTG
 25751 TTTTAAAAAG AAGGAAAGGG GGAAAGGCAG AAACCTCTGG AGACACGAGG
 25801 TCCCCTCTCC TCTATTTAA CTCTCGGCA TGGATTGTCT ATCTTGTCC
 25851 TTTCCCTCTC TTCCCAACTC CCAGTACCCC TTCTGGTGGC TGTGGCCAGA
 25901 AACTCAGCGA ACAGCACTTG TCCTATGCCC ATCAGTGTGC TGGAATTGAG
 25951 TGCACGGTAT CTCACCTGGC CTCAGCTCTT CGTCTCCAGA AAAAATAATG
 26001 GGCTGCCTGA GCTCTCCCCT CTCAGCTGAG TGGTGCCTGCT TTGTGGGTGC
 26051 CTCAACTCCC ACTTCCGTG GACATGCTTT CTTCCATGAG TATAAGAGCT
 26101 TCAGGTTACC ACCCGCACCC CCACTGGTAT CAGCCTGTGA CACCTCTGG
 26151 GCCTGTAGCC CAGGCCACA TCTAAAATA GAGGCCATC CCCCTCTGCT
 26201 ATAAAGCAAA GCCCTGAGAT TCAGCCTGCA AGGACTTACT GAGCACCTAC
 26251 TATGTACCTT GTTGCATCA CCCAGGATGC TGTGGACACA CCTCTAAATC
 26301 AGCCTCCTAC TGGGGAGATG GTTCAGAGGA AGAGAACCTT ACACGTAGTC
 26351 ACAGGGGATA GAAGTTAGGG GAACACAGGA GAGCAAAACA TTTCAAGGCAG
 26401 TGGGACCAGC ATGGACCAAA GCCCAAAGGA AAAAGGAAGT GTGGCCACCC
 26451 AGGGCATGGC AAGGGGCTGG AGAAGGCTGA GGTCAAGATGA CGGATGGGAC
 26501 TGCCAAGAGC CAAGGCCAAA AAGTGGCAGG ACCCAGCACT GGCAAGACTCC
 26551 ACTGTTGGGT CTGAGATTAT GTAGAGCAGG GTGGGGGTTG GGATTGTTCA
 26601 TGGTGTCTAG TAGGGGACAA GGGATGATTTC CTTACAGAGA CTCAGCAGCA
 26651 ACAAGAACTG GGCTCTCTAG TTTGACAGG ACCACCGAAG CCCCTCTGTA
 26701 CCCACTCTAGT CATTAGGCCA AGGGCCAGA GCCCTCTTAT GCTCTTGCCA
 26751 TTCTCTCAGA GCGGGCACCA GGGGCTAAAG AGAGTACCTT TTTTCCTTA
 26801 CAGGGGCCACCTTGGCCAC GGAATATGTT CAGCACATCC AACAGGCCCT
 26851 GGACATCCTC TCTGAGGAGG TAGGAGAGGG GTTACGTGTT CCTGGGTCCC
 26901 GCCAGCCACC TCCCTGGGAT GCATGTAGGC AGGCTGTGTT CAGTGAGATG
 26951 CTCACGGAGC AGAGACCCGC CATGAGTGA CACCTGGATG GCAGGGAGGG
 27001 AGGTGGCTGT CAAGCTCTC TGCAGGGAAA ATTCTCACTT GGCCAGAGAC
 27051 AGGGTTGTGT GGTAGCAATG AGCTTGCCTC TGAACCAATT GGCCAGGGTT
 27101 TGCGCCCAGC ACTGTGGCTT CGGGCAAGTG ACTTCCCTGT GTCTCAGTTT
 27151 CTCAACCTAT AAAGCGGGGC CACTCAAGAA GATTCACTGA GATACTACAA
 27201 GTTGCATCCC CTCTCTGGGC CTCAGTTCTC TCATTGGTAA AATTGAGGGG
 27251 AGTGGGAATT GGATTGTAGA TGACCCCCAG GTTCCCTCCCA GCAGTAGCCA
 27301 GTGCCCTAAC GAAACCACCC TCCACTCCCT GCAGCTCCCA AGGGCTTTCG
 27351 TCAACGTGGT GGAGGTCTAT GAGCTGGCTA GCCTGTACCA GGGCCAAGGC
 27401 GGGAAATGTG CCATGCTGGC AGCTCAGTAA GTGGACAGGT CACCGTCCCA
 27451 AGGCAAGGGC ACCTGGGGTG AGGAGGGCTT GCAGGTGCCA AAGGAGGAGA
 27501 CCAGTTGAGG CAGAGCCAGG CAGGCCCTGCC AGAGGGTAGA CATGGCTCAG
 27551 GGGCTGGAC AACATCAGGA AGTACCTCTA CATTGCAAA TGCCTACTGT
 27601 ATGCAAGGTG CCTCATTCTC CTGGACCCCC TTTTGCTTTT CTGTGAAAGG
 27651 AGACAGACCA AATGATCCTT AAGGCTCCCT GACATTGTCA GTGATTGCCA
 27701 AGGCAAACCT TGGCACGCTG CTCCCTGTG AGAAGCAGCA TGGGGCCATG
 27751 AGCTTTCAGG CTCAGCTCAT CCAAGCTTGA CTGTTCTGCC ATTAGGAGC
 27801 TATGGGACCT TGAACAAACC ACATAACCTC TATGAGCCTC TGTGTTCTC
 27851 ATCTGTAAAG TGGGGGTGAT GACACCTTCC CTGCAAGGTAGTGTGAGGT
 27901 CAAGAGGAAA CAAGGTACCT GGCCTACCAA GACTACCAAG AGCAGGGTCT
 27951 TAGGAAATAG CTCTTATTCC ATCCTTGATG GGGCTGTCC TTGATAGCTG
 28001 GGCTTGGAGG CAAGGTGCTA ATGGGCAAGA CAAGAACTCC TATCGGGGGC
 28051 TGGAAAGTCAT TAAAGCTCTT GAACCCCTGGT AGGAAGTTGC CATGTTCTGA
 28101 GGGCACAGGG CCTCCCCACAG TTTGAGTGTAT TATTGCTATG AGAGAGGAGG
 28151 TTCTCCAGGG AGCTGAGGAG TCCTACACCT GGGCTCAAAT GGATTTGCTG
 28201 CAAAGGTGAC CAGCTGGTTC CCATTCTTGC AGGGAACCCCT GCTCTCTCGC
 28251 AGGGAACGGC TCCCTCCAGAG TCTGTCTGTA TCGTGTCCCA TGGTGTCAAG
 28301 GTTGCCTCCA GCCGGTTGGC CCTCCCCACAG ACTTCCCACAG

FIGURE 3, page 9 of 33

00000000000000000000000000000000

28351 CTGGAGCCCT GAGGGAGGGT CCTAAGCAGT TGCAGGAAGA GCTGAGAGGC
28401 CCCCGGAAC T GAGGGAGCA TTCAAACCC AGGGACAGAG CCATCGTGGC
28451 TGGTTCTTA AACTCCAGTC TCCTGTCTAC CCAGTCTGC TCTGGAGAAA
28501 TCCCAGGGAC CACAGGCTTG GGAAGGAGGA AGGGGAATAG GC GTTCTGTC
28551 CACAGGGAGG TCCAGGCAAC AGCTTTCCCT CTTTCTCTAT GAACAATCAT
28601 CCTCTGGACC TCAGGGCTCC TGAGTTAGCA TTCTGTAACC TGGGTCCAAG
28651 AATCAGCCAA AGGTGTATTG TGGGGATACT TGTGTGTCAC CCCCCGCCCT
28701 AGGTAAGGCA GCACAGGCTG CAGGCCCTG GGGTAGTGGC CTGCTCTGTG
28751 TGTAGAGGCC AGCCTCCCAG GAGGACAGAG CCACAGTGCC CCAGGGAGCC
28801 TCAATACAAAC ACTCCCTGTC TCACAGGAAC AACTGCACCT GCCTCAGACA
28851 CTCGCAAAGC TCCCTGGAGA AGCAAGAAGT GAAGAAAGTG AACTGGAACC
28901 TCCAGGTAAG CCCTGCAGCC CTTCTCTTAC TGACCCAGCT GGGGGGGCCC
28951 CTGTAATCTCA AGGACTGGGA AATCGAATGC CCAGCAGGAT GTGGCCAAGA
29001 GCAAGCCACT CCCTAAAAGC AGATTGCAGC CCCTGAAATA CTTACCCCTG
29051 CAAATTGAAC ACCAAGGCCA GGGAAAGGGAG TGAGAGACCC CAAAGTGGAA
29101 GCTGAGAAAAA TCCCCTCTC CCAGCGGGTA GGCAGCAAGA GATTCCCAGA
29151 GTAGACTCCT TGTGGTAGGG CCCATCCCC ACCCAGAGCC ATGTGTAATA
29201 ATTACTACTC ACTTCCCTCCC CTCCCTTCAT TAAAAACAAA AGGCTTAGGC
29251 CCGACACAAAT GGCTCACGTC TGGTGTCCCA GCTACTCAGG AGGCTGAGAT
29301 GGGAGGACAG CTTGAGGCCA GGAGTTGGAG GCTGCTGTTA GCTATGATGA
29351 TGCCATTGTA CTCTGCCTAG ACAACAGCGT GAGACCTAT CTCAAAAAAA
29401 AAAAAAAGAA AAAAAGAAAA AGGCTTAGCC CTGCCCTACT TAACTCTACC
29451 TCAAATTCTC CTTGCCCTCT CTCTGCCCTC TTCCATCTCC CCACCTCCAC
29501 TCCTGCTTAT GTCTCTGCCT CTATTGTTCC CTCTCAGGCT CAGGTAGCAT
29551 TTCCATTCTG CAAACTGACC CTCCCTCATT CACAAGGCAA GTCTGCTTCC
29601 CTCCCTAAG GAGCTTCCCC TGCTGAACCT TCACCCCGGG ACATCTCCCC
29651 ATATCACATT CAGTCTGTAC TTGATGGGC CTAAAAGCCC CAAAGGGTTC
29701 TCATGTTTC ACATCTTGGC TCATTTTTCC AGATGGATGA TAAACTCCTT
29751 GAAGATAAGT ACATCTAGTC TGTTCTTTT ACATTCCATG CTTGGGTACT
29801 TAAATCCAGC CACCGTGGAC TCTCCTCCCC CAAAGTTCAT GGGCATTTTG
29851 GGAGCTGGTG TTGAGATGCT CCCCCATCTGA CCTGCAGCCC CATGTTCTAA
29901 TTGACCTCTT CGTCAGTGA GAGGGAGGGGA GGACTTGGC CTATGCAATC
29951 TGGTCAGTGG CTCAGACCCA GCCTTCAGG CAGAGCTTT GGAATGGGAC
30001 TGGGTGGAGC TGTGTAGCTA GGGAGCTTCT CCCACAGGA GCCGCTGGGT
30051 TCAACTCATC TCTGATCCTG AGAACCCAGCA TAGGGCTTTG AAATGTCCGT
30101 GCCCCATGAAT GGGTGGAGAA TAAAAGTATG TTTGCATCCC ACTAGAGTAG
30151 CCCCTTAAAG TCACTGTCCT TTAGGGTAGAG TTGACTCCCG TCAACAACCA
30201 ATCCAAGGCA GCAGGACTGG ACCCTGTCTG TGCAGCCTTG CCAGGAGGGT
30251 TGAGCAGCTT CTCTCTCTGT CCCCCAGCATG GCATCTCCAG TTTCTCCTAC
30301 TGGCACCAAT ACACACAGCG TGAGGACTTT GCGGTTGTGG TGCA GCTTT
30351 CTTCCAAAC ACACTCACCC CACTGAACCGA GGTGAGCTGC AGGTATTTA
30401 GGGAGGCTCA CGTATGGGGG CCTTATCACA GACGATGGAT GTATTTCTT
30451 CTCTAAGTGG GCTTTTTTTT TTTTTTAACC ATCTCTCTCC AAGAGGATTC
30501 CTGAGGGTGG CTTTTCCAC ATTACCTCT TTTTGTGGGG GCTGGCTGT
30551 GATTGGAAC T CAGATGTAATCT TGAAAGGAA ATCAATAGTG ACTAAGCTCC
30601 CAGGCCTGGC CCTGATGTTT TCTGGATTGG GATAGAATGG AAAGCTTCC
30651 AAAATGTTA CTCTTTCAA CTCTTAGGAT AGGGGTGCTG AAAGAAAAGG
30701 GAGAGACTAT GGGTGGGTCC AATTCTGTG TGTTAAAAA GAAAATTCCG
30751 GCCGGGTGCA GTGGCTCATG CCTGTAATCT CAGCCTTTGG GAAGCCAAGG
30801 CGGTTGAATC ACAGGAGTTAG GAGTTTGAGA CCAGCCTGGC CAACATGGTG
30851 AAACCCCGTT TCTACTAAAA ATACAAAAG TTAGCTGGGC GTGGTGGCAG
30901 GCACCTGTAA TCCCAGGTAC TCGGGAGGCA GAAGITGCAG TGAGCTGAGA
30951 TTATGCCACT GCACTCCAGC CTGGCTGACA GTGCGAAACT CCGTCTCAA
31001 AAAAAGAAA AAAAAGAAAA AGAAAATCTA AATTCTGGGA GTTTTCCAT
31051 CAGTATCTGA GCAAGTTGGC AGGAAAGTTG AAAGAATGAA AGGAGACATG
31101 CCCAGGGCAC CTGCTGGGAG AGTGAAGTGGG GCTCAGGTAG CAGAGCCCTT
31151 TCCAGGGATG ATAACCTCTT TGCCGTTGGT TGCA GAGAGAGGG GGACACTGAC
31201 CTCACCTCTTCT TCTCGAGGA CTGTTTCAC TTCTCAGACC GCGGGCATGC
31251 CGAGATGGCC ATCGCACTCT GGAACAAACAT GGTGAGCAGC CAAGGGCCTG
31301 GTGGGCCTTG TCAAGGGGGG ATCTAAGGAT ATTGACACTC TGTCTCACAA
31351 TGGCAAAACT ACTGGAGACA TGGCTCCTT CTCCCCAAAG CCCAAAGTGG
31401 CAGCACACCT TATGGTCTT GATAGATTAAT TCCAAAGGG AAAATACCC
31451 ATATTATCC AACACCCTT GAAAGTTATA CAAACACACA CTCACACAAAC

450
400
350
300
250
200
150

31501 TTTATTCTT GTTCCTTCAG CAATGCCAG GTACTGCGAG GGGATCCCTT
31551 TGTAATCAGA TAGTTGGCT AGATGAAAAT ACCAACTTCT ACCTCGTACT
31601 GTGTGACCTT GGGCAAACGA TCTCTCTGGC CACCTGTATC AACATCTATA
31651 AAACAGTGAA ACAAGACAG GTCTCAGACA ACGCATTGAG ATCATGTGTA
31701 CATGGCACCT AGCACAATAG TTAGCACTCA GCAAATGTCA CCACCATCG
31751 CCTTCCAAGC ACTCAGGGCT CAACTCATAC CCAACTCATT TCTCTAAACA
31801 TCGAAAAGTG GAGATCCACA CAGCCTGTTT TCCGAGGCTG ATACCTATT
31851 CAGTCCTTTC TGATGGGAAG AAGGGACCTT ATGAAATGAA CATACTGTCT
31901 GGGGTCTTT CAGGGACACC TGCCTGGTGC TTCCACTCTG CCTTCTGTGG
31951 CTGGCCACCA GCAACTGAAC GGTTCCGCA CAGCACTTGA CCTGTCACCC
32001 CCAACAACCTG GATCCTTTC CACGGAGCAA ATGAAATGCC TTCCCAACCC
32051 AATGGTTCT TTTAATCCAG GCTCAGTGGG TAACACAATC CCCACCCCCA
32101 CCTGTATGTT CCCTCTTTG TCCTATGACA ACTAAACAAG CTACATTCCA
32151 GCTCCTTTTA TCACAGTTTC AGGCCCCGTAG TGTCTTGCC AACCAACCGCT
32201 GTGCAAACGT TCCCACCCCT GTCAGCTCAT CCAGTATGTC CAGCATCCCA
32251 CTCGGCTGAC TCACAATATT GACTTCTCC TTAGCTATAC CATCTCCTCC
32301 TCTCTAGCAA CCTCTCTTT TAAGAACAGC ATGTAACACTG GCTTTATCCT
32351 TGGCCTAGTT AATGGCAGAC TCAGCTTATG TCGACTTCCA TTGTCAGGGG
32401 GTTTCTCC TGTGGACATC ACGTACCTGC CCACTCCAAG AACCTCTATT
32451 GTACTCTTC AGCCCAAGAC TCCGGATTGT AACCAAATG TCTTGAGTT
32501 TGCCCCACGA TTTTAAATC AGTTTATATG GTATAATTCC TGTTCCTTT
32551 GTGGATTTCG TTTGAAGGC GGTTGTCTT CACTGGCTGA ATCATGTGAC
32601 TTTATTCTT TGAAAATC TTCCCAAAGA AAGGGTACCT ATTCCCTGTT
32651 CTTTTCCCC TGAGACCTCA GGGGATTCCA CAGATGCCCT TGGCCCTTCC
32701 TTCCCAGTT TTTCATCAAG GTATGGCCTT CCTACCAAGGT GGCACTCCA
32751 GTCTGCTAA ATCTGGGACCT CTCAGGAAT CTCCTGGGGC TGGATAGCCA
32801 TAGTGACCGC TGGACATGA AAAAGAGTCC ATTGGTTCT TTTCTTGTGA
32851 ATTAACAATG TAGCTCTGGC CAGGCACGGT GGCTCATGCC TGAAATCCCA
32901 GCACTTTGGG AGGCGAGGC AGGTGGATCG CTTGAGCCCA GGAATTAGAC
32951 ACCAACCTGG GCAACACAGG GGAGATTCTG TCTCTACAAA AATAATCAA
33001 ATATTAGCA GGTGGTGGG TGCATGCCCT TAGTCCCAGC TGCTCAGAAG
33051 GCTGACGTGA GAAGATCACT TGAGCATGGG AGGTCAAGGC TGCAATGAGC
33101 CGAGATGGCA CCACCGCACT CCAGCCTGGG CAATAGAGTG AGACCCCTATA
33151 TCTCAAAAAAA CAAATAGAAA AAAAAAATAT ATGTAGCTCT GGCCCTCTCT
33201 TCTAAACCGAG TTCAGTAGCT CTTCCCATTC ACCCAGGTA GAGGCCTTTA
33251 TTTCATAAAG ATAAGTGGG GGAGTTAGA TATGAAAACA AAACGTAAAC
33301 ACCGCACTGG AGCTATTGTG GAAACAAAAC AAGACTGTCC ATGGTTCCCC
33351 AGCCATTATT ATCTCAGCCA TACCCCGAAT TTCAAAATAA CAAAACAAA
33401 ACTAAACGCCA TCCAGGGGTT TCTTATCCTA GGCTCTATAA TTTGGGTAAA
33451 TAATTATACA GTCTAATGTT TTCATCCAAA GCCAATCTTA GACATAAAGC
33501 TGTAGCATGA TGCCAACTTT TCAGATCGGC TTCTGGCTGG AATTCACCC
33551 CTAGAGTAAC AAAAAAATAA TAATAGACCA TTAGAGCTGG AACAGACTGA
33601 GAGGTCACT AGCCAGAACAA TTCTGTAAC AAAGCATAGA AACATGAAGC
33651 AGTTTGCCTA ACATAACACA GACTGTTCAT GGCACAAGGG GGATTACAGA
33701 CCAGGTTTT CTAGCTCTT CCTGGTGACC TGGGCATGCC ACCACCCCTCC
33751 CCACTGCTCC CAACCTGATA AGCACATATA TACCCGGTGA ATTCTATGCT
33801 CACAATTAGA GTCCTATGAC ATAGTGTCTG CAGGCTTTGG CTGATGTTCC
33851 CATAGTGTCT GCAGGTTTG GCTGATGTTCC CCAGGGTTCC CTACTAGGA
33901 GCAAAAGCA CCTTAAACTA TTTCATCTTA TTTCATCTCC TGCCCTCTCT
33951 CTCACGTCT TCTCGAGACT TTTGCAAAGG CAAAGCCAGA AGCTCCAGCA
34001 GCACCAAGGGG ATATTTCTT CTTCTCTGC CTTCTCTGT CTCTTATCT
34051 GAAGAAGTTT CTCTTCTCG AGGGCTAGTC CTCTACTGCT GCCTCTACTC
34101 CCTCTTCTGC AGAAATCTG CTCTCAGCCA GTGTTGTAT CTCCCCAGGT
34151 GCTGGGTGAC AGCTCAGGC TCCTAACTGA CATCCCTGTC TTCAAGCTTA
34201 GAGCTCTTAG AATCGTGAAT CTCAGCTCTG GCTGCATATT AGAATCATT
34251 AGGGACATTG TGTATGTGT TATGTATGT TATGTATGT TGTATGTGT
34301 TGTATGTGT TGTATGTATG TATGTATGT TATGTATGT TGTATGTATG
34351 TATGTATGAC AGAGTCTCAC TCTGTTGCC AGGTTGGAGA GCAATGGCAC
34401 CATCTCAGTT CACTGCAACC TCCGTCTCTT GGATTCAGC GATTCTCCTG
34451 CCTCAGTCTC CCAAGTAGCT GGGGTTATAG GTGCATGCCA CCATGACCAG
34501 CTAATTTTG TATTTTTAGT AGAGACAAGG TTTCGCCATG TTGGCCAGGC
34551 CAGGCTGGTC TTTAACTCTT GACCTCAGGA GATCCACCCA CCTCGGCCTC
34601 CCAAAGTGT GGGATTATAG GTGTGAGCCA CAATGCTTGG CCATCCAGGG

FIGURE 3, page 11 of 33

34651 ACTTTTAAAAA CAATTAGTGC CTACAGCCAC TTTGGAAAAT TCTTTGGTAT
 34701 ATTTAATAAT GCTGAACACA TGTATTCCT GTGATCCAAG GATTCACTC
 34751 CTAGGTATAT CCCAAAAGA AAGGTATATA TGTGTCTACC AAAAGATACA
 34801 CACAAAATAA TTCACAGCAG CACTATTTAT AAATAGCCCC AACTAGAAA
 34851 CTCCCCAAAT GTCCACTGAC AGTAGGATGG GTGAGTAAAC TGTGGCACAT
 34901 TCATACTAACG GAATACCATA CAGGTCTGAC CGCATCTGTG GCTTAAACA
 34951 AAAATCAAGC AGGGTGATGT GACACAGAGT AATGGCTGGG AAGAGGGAGG
 35001 CCTCACTGAA GAAGTGACAG CTGAACAAAC TTCAACAAAC TACAATAATA
 35051 TCTATAAAGT TCAAAGCAA GCAGCTTGGC ATATGGGTT AGACGTCAGC
 35101 ATGATGGTGT AGAGACTCAC TGGGGGATGA ATAGTCTGG AAGAAGGTGG
 35151 AAAGGGGCTT TTGAGGACTA TAATAGTCTG TTGCGCTGACT GGATGCTGGT
 35201 ATGTTCATTT TATCGAAACT TATCTGTTGC TCACTTATGA TTTGTACTCG
 35251 TTTCTATGT TATGTTAGCT TCAATTAAA GTTACTTGA GGCGGGGTAC
 35301 AGTGGCTCAC ACCTGTAATC CCAGCACTTT GGGAGGCCGA GGCAGGCAGA
 35351 TCCCCTGAGG TCAGGAGTTC AATACCAGCC TAGCCAACAT GATGAAACCC
 35401 CATCTCTACT AAAAATACAA ATTAGCCAA GCGTGGTGGC ACGTGCCTAT
 35451 AATTCCAGCT ACTTGGGAGG CTGAGACAGG AAAATCGTT GAAACCAGGA
 35501 GGCAGGGGTT GCAGTGAGCC AAGATTGCAT CATTGCACTC CAGCCTGGGT
 35551 GACAAGAGTA AAACCTCTGTC TCAAATTAA AAAAAAAA AAAAAAAAGT
 35601 TTACTTGAAA AACAAATATCA GTGCCTGACC GGGCTTATCC CCAGAGAGTC
 35651 TGACTTAATT GGTCTGGAGT GCGAGCTGGA TTCGGTACTT TGTAAAGCT
 35701 CCTGAGAGTA TTTTAATGTG CAGGGTTAT GAACCGCTGC CTTAGATCTG
 35751 GTCCCCACAG AGAAATCAAG TAATCTGTAT AAAAGAAAAC CTGACCCAGT
 35801 CACTCCCCTG CTTTCAAAC TCCAAAGCCT CCCACCTCTG AAGGAGGCCAG
 35851 GCCAGGCCCA ATAGCACAGC ACACTAGGCC TCTGGGACTT GGCTGGTTC
 35901 ACCTGATTAAC CCTCTCTGGC TACCATTTCC ACCAGCGTCT GCCTCGCATG
 35951 TTACAGTCTA GTGACTCCAG CAGCGTCCTG CACCACCTGT GGTGTTCCAC
 36001 ACCTCTGCTA ACTCTTGCTC TCCTCCTTCT CCTGGATTGC CCTTCTCACC
 36051 TCCCTGCCCA CTCCACCACT CAACTCAGGT GCCACCTCCT GCAGGAAGCT
 36101 ACCTCTGAAT CTCCAGGACA GGCCAGTGGC CCACCCAGGT CCATTACACC
 36151 CTGCCCAGTC CTGTCATTG CTACGTGGTT GGTAGCCACA GTGCCTGGCT
 36201 TAGGAAAGAC TGGTTCTAGG AAAAACAAATT TCATTCCCTG TGCCAGCTC
 36251 CAAGCCTTCC CCCGCCAACG TTCTCCATTG AGGTCTCTGT GAATTAAATT
 36301 AATTCATCCA TCCATCAAAC AAGTATTAC TGAGCACTAA TATGTGCTAG
 36351 GTACTGCTCC AGGTGCTGAG GACTCAGCAG TGAAAAGATG ACTGCTACTC
 36401 TCATGGGACA TACAGGATAG TAGGGAAAAG ACAGATAATC AACAAAGGTCA
 36451 TTTCTGACCA CATCTGTGGT TTAAGAAAAA GTCAAGCAGA GTGATGTGAT
 36501 ACAGAGTAAT GGTGGGGAG AGGGAGGCCT CCCTGAAGAA GTGACAGTGA
 36551 ATTGAGAAGC GCATGTCAG GGGTTGCCAG GCAGAGGAAA TAGGACCCAC
 36601 ATGGGCCTAG AGTCAGGAGT GAGCTTGAAG TGTCTGAGGA ACTTAAAGGC
 36651 CAATGTGACC AGAGGGAAGT GAACAAGGTG AAAAAGTTGG GCAGGGGCCA
 36701 GGTCCCTAGA TGCTCTAAG CAGTAGAGTG ATATGCTCTG GCTTACCCCT
 36751 GGGTCCGTGT ACCCTGGACT GGAAGAAAGC AAGGGTGGAC CTGGAAAGAC
 36801 CACTAGGAGG CTGCTGTTGA TGGGTGAGAG AGGAAGGGGG CTGAGAGTAG
 36851 GGTCAGGGCA GAGGAGGAGA GACGCTGTCG TGGGCTGGCG GATGGATGAT
 36901 GGGGAAGAGG AACAAAGGAT GACTTTTTGG TTTGGGGTCT AAGAAACTGG
 36951 GTGGATGATT GAGCAGGTAG AGAAAAAAATC AGCGTGGAG GAAAAAAAT
 37001 CAAGACTTCT GTTTGACCA TGGTCAAAC TGCCTTCCAG ACATCCACAT
 37051 AGAGGTATCA GGATACAGAA GTTTGGAACACT CACAGAGGAA GTCAAGGCTG
 37101 GAGATTGAAA AAAAAAAA AAAAAAAA AGTGGGGTTA TTAGCATAGA
 37151 GGGCCAATAT GGTGAAACCC TGTCTCTACT GAAAATACAA AAATTATCCA
 37201 GGCATGGTGG CATGCACCTG TAATCCCGC TACTCAGGGA GGCTGAGGCA
 37251 GGAGAATTGC TTGAAACCCAG AGATGGGGTG GAGGTTGAG TAAGCTGAGA
 37301 TCGTGCCTTGC CACTCCAGC CTGGGTGACA GGGCAAGATT CCATCTAAAA
 37351 AAAAAAAAG CCACTACAGG ATCAACTAAG AGCTCCTAGA GAAAGAATAG
 37401 GTAGGTAGAA AAGAGTGTAA GGCGCAACTAC CTAGCCCTGG GCATTCTTIC
 37451 CAGCTTCAA CTCCAGTGGAG AGATGAGAAG GAGAGTGTGG AGGTAGATGG
 37501 GAAATGAGAA ACAATGCTGT GTCCAGAGAG CTAAGAGAAG TCAGTGTTC
 37551 AAGAGAGACA GAGCTGTCAA CTTTGATGGA TGCTTCTGAG AAGCCAAGCA
 37601 AGTTGAAGAC AAAAAAAA AAAATGATCT TTGGCTCTGC CCATATGGCG
 37651 ATCGTTGGTG GCCAGGGCCA GAGCTTCCAT CCAGCGATGG AGACTGCAGA
 37701 CTGGCTGGAG CGAGCAGCAG AGAGAAGGAG AGATTAGGAA GTGCTGCCAG
 37751 CACCTATAGA CAGCTCTTCC CAGAAGTTAT GAGAAGTAAC AGCCACGGTC

37801 ACTGGAGGGG ACATGGATCA AAGAAAGGGC AGGTGAAGGA GGGGAGATGT
 37851 CGGAGCAGGT TGTGACTGA CGAGAAGGAA CCAGTAGAAA GGGAGAAACT
 37901 GATGCACTCA TCAAACCCCTT GTAATCACGA TCATCTCTG TGTGAATTAG
 37951 TTCTGGGTTCTGGGATAGC ATCGGGAAATC AGCCGCGCTG ACCTTTAGCA
 38001 TTTATTCTGT CACTGTTACG ATAGACTTGA GTTTCCTCAG TTCTTAAGAA
 38051 AGTGGAAATA ATACTACCTT ATGTATGTAA GCCACCCAA TCACACGTGC
 38101 TTTCATGCCA TCTTCTCATT TGATGCTCAC AACAAACCAA GGTGTCAGGA
 38151 CAGGATGTCA TGCCGTCCC TAACCTCAGA GAAGTAGCCA GCCCAGCACT
 38201 GCACAGCTCG TTACCAGCAG AGCCTGGATC CCAGCCATCT GCCCATCGTG
 38251 CTCAGTCTAG TCACCCCTAGC ATCTCTCCCA GGAACAGAAC TGCCCTCCCT
 38301 CCTCCAATTG TGTTACTAAG GAACGGGTGT AAAAGGCCTT CAGACAACAG
 38351 AAAGTGAGAT CATGGGCCAG GTGTGGTCAT GAGCTCAGAT AGTGAACCTT
 38401 CACCTCCCTC CCTGGCAATA CCCTGTGGTC AGGAGCAGGC AGATTACACA
 38451 AAGAGTGGAG GCTAGACGTT CCAAACAGAC TCTGAATAGG TGACAGTGCC
 38501 AGGGGCTCAT TCTTCTCAGT GCTGGCCACA GGTTGGCCT GGCTGCTGGC
 38551 TAAAAGGTGC CGGGGAGGGG GATACAGCAG CTCCCAGCTC ATCCTCAGAG
 38601 GGTCCCTGGGA TCAAAGGTAT TTACACCCAG GGATATTCA GATAAATCTT
 38651 TTCATCTATG TGGAAAACAT ACAAAAGTGGC GCAAGTGAGA AACTCCGATT
 38701 TCCTAAGGTT GACAAGTCAA GTGCAGTAAT GATGTCATGG TAACCAATAT
 38751 GTTTCCAAAC TTTCTTAAGG TTGACTAGCC CCATGCACTT TGAGAAGTTG
 38801 GTAAATAGGA TTGTCGTCGT TTTATAAAAT TGAAAACACG GTGCTTGCA
 38851 ATCACAGCCA CTCACAAAGG AAGCCAGAGA TGGTCCCAGC CCCTCCGCAG
 38901 ACTTCCTGTG GACTCAGGAC TGGTGGTCTC TCCTGGCCT TGCTGTACCC
 38951 GGCAAATCCA GGGGCACAGA CTCAGGGTTC TGCCCTGCCG ACAGATGCTG
 39001 CCTAGCCTTC TGTGTGTAT AAGTCAACTC CCGCTCAGCC CCAGGCTGCT
 39051 GGGTCCCTGC TGTGGGCCAA AAACCAGCCA CTTCGCTGGT TTCTATCCCC
 39101 CACCCCGTTC CCGAGGGAGG GGCTCTGGTG TGAGACACCC CCTCAGAGAG
 39151 GAAAGTGCT CCCAGCTTTG GAGAGAATCG AGGTGTCTT TCTCTCTC
 39201 CAGCTGGAAC CAGTGGGCCG CAAGACTACC TCCAACAAC TCAACCCACAG
 39251 CCGAGCCAAA CTCAAGTGC CCTCTCCTGT GAGTAAACGT CCTGCCTGCC
 39301 CCAGGTGAA CAGATGCCCTG GGGTGGGGGT TGTCTGTCTC CCTGGAAGCA
 39351 CAGAGGAGTC CCCGGGGATG CTCCCTCAAA TGCGGCTTCA CTCACTGCCG
 39401 TCTTCTCAAA TCCCACCTGT CCCAGTGCC ACGGAAACTT CTCAGTGTGT
 39451 GGGCAGCCAT GGAGGGAGGG GAGAGGACGT TCAACAGCTC CAACCGAAGG
 39501 GAGGACAGTC GCTCAGGGAG GCAGATGAGC ACTGGGGGT GTCTCGGGTC
 39551 ACCCATTCTC TCCGAAAGCT CTGATGCATC CTCAGTCITA AAAGTGCACC
 39601 AAGGCCAGGC GTGGTGGCTC ACGCCTATAA TCCCAGCACT TTGGGAGGCC
 39651 GAGGGCGTCA GATCACCTGA GGTCTGGAGT ACAAGACAG CCTGGCCAAC
 39701 ATAGTGAAC CCCGCTCTA GTAAAAATCC AAAAGTCAGC CGGGCATGGT
 39751 GGCAGGGCGCC TGTCATCCCA GCTACTCAGG AGGCTGAGGC AGGAGAATCA
 39801 CTTGAACCCG GGAGGCAGAG GTTGCAGTGA GCTGAGGTCA TGCCACTGCA
 39851 CTCCAGACTG AGTACAGAG TGAGACTGTC TCAAAAATAA AGTGCATCAA
 39901 GCAGCTGTCC CGTGCCAGGC AGTATACTAG GATCTGGGA TCGGGAGGCA
 39951 AAGATAAAAT AGACTCAGTG TCTGTTCTG GAGCCTGCAA TGGTCTTCCT
 40001 CCCTGCCAC ACCCACTGCC CTTGCCTGGC CCACCTTCGA AGCCTGTGAC
 40051 TTGTCTCCCC AGCTCTCCTC TCCCTCTTCT CCATCCACCC TACACTTGCT
 40101 GCCAGACACA GATAGACCTT CCTGGAAATA ACTTGCCCCA TCAAGGCTGC
 40151 TTGAAATCCT TGCTGTGATCC CTACTGCCA TTGACCCAGAG TCTGGAGGG
 40201 GGGTCACCTC CCTCCATGAT ACACACTGCA CTCCCTGCCG GTGGGATCCAT
 40251 CTCCCAGGAA GCCCCACGAC TGCCCGCATC CAGGCCCTTTC CTTTTGCCAT
 40301 CTGTTCTGG AGGTTCATCT TCCATCTGCT ATGAGAACAT CGGCCCTCCCT
 40351 CCAGGTCCAG ATGTTGCCCT TACTAAGCGA TGGTTTCACC GTCTCTTACC
 40401 TACCATTCCT GTCTCCAGAC ACTGACCCAT GTGGGCTCC TTTCTATT
 40451 GTACCTCTCA TGAGACACCG ACCCAGTCTC CTTTATGATG TGATTGTT
 40501 TGCACTCTC AACTCCCTCC TGGGCCACAA GAAAAGATGT CACATCTAA
 40551 CCCTCCAGTC TCATCACAGC TTCCAGCAAG GGGGCTAAAC ACAGCACGTG
 40601 CCCAATTCAAC ATTCACTGAG AGGAGAGTGG AGAGGGGCAT AGGAAGGCAA
 40651 GAACGCACAC GATCTGCCA CATGCCTCCC CTCCCCGCCCT TTCTGATTTG
 40701 GGGATCTTCA ATCTACTACA AAACCAGCTG TCCTTCCATG CTGCCCTTCC
 40751 CTGATTCTG GGTAGTCCCTG GGATGGGAGA ATGGGGACAG TTGTGACCAC
 40801 GAGGAAGCAG AGGTGGGAGT TCTACAGGCC CCACAGGGCT CTCTGCCATT
 40851 GGTCACTTAC CAGTTCCCAA TCTTTCAAAA TCAGGTTGA TGGCCAAGGA
 40901 AACGCTGGTG AGAAACCAAA AGAAGGTTCT AGCTGGGTGT TGACCTCTTT

FIGURE 3, page 13 of 33

40951 AGAGGCCAT CCCGCTAAAG AGGGTTGGG CACAGCCTAA ATGAGGGAGC
 41001 TTTACAAAAG GGAAGCTCTG TGAAAACGTG CAGGGTTATC GCAGCATCTC
 41051 AGGAATGGGG ACTAGGCAAG TCTTGGCTTG GTGATGGATG GTTCACGGAG
 41101 ATCCCTTCCA CTGACCCCCG CTCCTCCTCC ACAGGAGAGC CCTTACCTCT
 41151 ACACCCCTGCG GAACAGCCGA TTGCTCCAG ACCAGGCTGA AGAAGCCCC
 41201 GAGGTGCCT ACTGGGCTGT CCCAGTGGCA GCGGGAGTCG GCCTTGTGGT
 41251 GGGCATCATC GGGACAGTGG TCTGGAGGT CAGGAGAGGT GGCGGGAGGG
 41301 AAGATCCTCC AATGAGCCTG CGCACTGTGG CCCTCTAGGC CGGGGGGTGG
 41351 GTCCTCACCC TAAACTCCCT ATAGCCACTC TCTTCACCAG CCTCTGCC
 41401 AGCCACTCCC GGCCACCAAGG ACATGCTTCA ATGCCTGGTG CCATAGGAAG
 41451 CCCAGGGGAC AGTCACAAC TCTTGGGGC TGGGCTCTT CCAGGCCTAT
 41501 GCTCCTGGAA TGGATACATT TAAATAAAAGT CCAAAGCTAT TTTATTCC
 41551 GGTTTGCGCT CGTGAAGCAC TCACCTTCCA TCTCTGTGC AGCCCAGGTG
 41601 TGGGAGCTGC CACTTTTGT GCCCTGCCTC CAGCAGGGCT GCCCAAGCCA
 41651 CGACCAACCA GAGCCCAAAC TGCCCTGCCAC CACGAGCATA TCCTCAAGTC
 41701 ACCAAACCCA CTATTTCAAA GGCAGAAAAA ATGCTGGTCA CCAGGTGGT
 41751 GCTGGAATT TGGAGCTGGC TGGTTGCCAT TCAGTCCAAT CCAACACATA
 41801 CCTATTAAAGC AACTGTTTTG TATCCAGGAC AATGCGAAGC ACTGAGGTGC
 41851 CTCCTAGGCT GTGCATGTG CAGCCTGGCA GAGAGGTCAA ACTCCTTCAA
 41901 TAACCAAGAA GCCACGTGAT GATGTGTAAC TACTAGGGCA TCAGTAGGTA
 41951 AATGTGTCCTG ATTGTTTAA AGAATAGAAA GGGTTCTTCG GGGAAAGTTT
 42001 CTTGGGGGAG AGCAACCTTC ACATGTCATT TTGGGAAAAG GAATAAAAAAA
 42051 TGATTGGGAC ACAAAATACCT CCTATATTCT CAACCTGATT TTCTCAAGGT
 42101 GCTAAATTAA GGAAAAAAATT CCTATTCTA TATGCCAGG TTTCTGAGGG
 42151 AAAACTAGAG AGAGTCTGAA AATATGGGCT GCATTCACTG AGCCCCTGCT
 42201 AGGGGCGAGG CCCCGTGCCTG GAGGCCTTCC ACAGATGGTC TCTTTATGC
 42251 TGCACAAAG CCCAGGGAGG GGGTAAAAGG AAAATCTTG AAAATAGAAG
 42301 TGATGCTTGC GCAACACCGT GAATGTACTA AACGCCGCA ATTGTTCCAT
 42351 TTAAAATGAT TAATGTGTA TCATGTGAAT TTCACTTCAA TAAAAAAGAA
 42401 TCCAGGGAGG TAGACATCAT CTGCATTGTA AACCTCTCTC TGATCCTGAA
 42451 GTCCGGGATG ATAAGAGGCC TGAGTCACAA TCCGGATGC AACACTGAAA
 42501 TGCTGTGCC TGAAGCTGCC TTGCCCAGCC TGAGCCCAGT GTCCCAGGCT
 42551 CTGCATCTGT AAAACTGGAA GTAAGAGTAC ACATTTGCT TATCTCACGG
 42601 CGCTGCTGAA AAATAAGGAA CCGTGTGTGA ACCTCTAACT CTAAAATGCT
 42651 GCACAACTGA AAATGCCCTT TTCCCTCGGT GAAAGTTGG GATAAGGCC
 42701 AGACTGTTGG GGAAGATGTG AGACCCAGAG ATGAGTTGG GAAAATGGGG
 42751 TAATAACATA TGGTGGAGA GTGCCCGCCT TCCTCTCAGG GAGGTTCATC
 42801 ACCTTATCTC TTTCTGTCAC AACAGAGAAC CCGGAGGACC TATAACCCAGT
 42851 TCCGTGTCT TCTGGGCTTC AGTGTCTGTT TCTATACAAT GGGAACAGCA
 42901 TGCATTCCCC TGCTTTTCC TATAGACTGG AAAACGTGGT GACCAAGTC
 42951 CACATCCCAG CTTATGCTCC CGGCTTAAGA CAGTGTAAACG ACAAAAGGTAA
 43001 CCCTTACACT CCTGGTTTGA GACAGTATAA CGACAAAGGT AACATAGGAA
 43051 GTCAAGGAGT TCGCTTCACC GCCCCTCCCC CCACCCACC CTTTTTTTTT
 43101 CCTGCAAGTT TCTATTCTTC CCGCAGCTCC TACCTCAAAG CAGCATGGAT
 43151 TCATAACCAC AGGCTCCCCT CATTAGGGCT TGGGGAGGGA GGGTGTGGA
 43201 ATCCACACTG CCAGAGTAAT CCAGACTAAA ACATCAACAA ATGGTCCCAG
 43251 CTGGTTCACC AAGGAACACT TGGCAAAACA AAGAAATCCT GTCTGGAGCG
 43301 ACACGGACAC AGCCACAAAC CAGTCACCAA ATTCCAGCA AGTATGTGCT
 43351 AAGAAGCCAA AAATTAAGAA TACGTGAGAA GCACCCACTT GAAATTGGTG
 43401 GTATTACATA CATACACTGG CTGTGCGCCA TGGGGTTTT CTGTAGGAAA
 43451 ATGTCCAGTC TAGCTAGAAC GGCACCCACA GCCACACCAT GAGCAAAGCC
 43501 ACCAAATACC TGAGGAAGCC ACAGTCCATG GCACTCCCCA TGG

FEATURES:

Start: 3000
 Exon: 3000-3098
 Intron: 3099-4565
 Exon: 4566-4637
 Intron: 4638-12578
 Exon: 12579-12644
 Intron: 12645-16277
 Exon: 16278-16338
 Intron: 16339-16891

Exon: 16892-16950
 Intron: 16951-18968
 Exon: 18969-19036
 Intron: 19037-19303
 Exon: 19304-19347
 Intron: 19348-19582
 Exon: 19583-19682
 Intron: 19683-23297
 Exon: 23298-23349
 Intron: 23350-24668
 Exon: 24669-24749
 Intron: 24750-26803
 Exon: 26804-26869
 Intron: 26870-27334
 Exon: 27335-27426
 Intron: 27427-28826
 Exon: 28827-28905
 Intron: 28906-30276
 Exon: 30277-30381
 Intron: 30382-31185
 Exon: 31186-31281
 Intron: 31282-39203
 Exon: 39204-39278
 Intron: 39279-41134
 Exon: 41135-41335
 Stop: 41336

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
1573	A	T	Beyond ORF(5')			
1737	T	G	Beyond ORF(5')			
2498	T	C	Beyond ORF(5')			
3097	C	T	Exon	33	A	V
3116	A	G	Intron			
4823	G	A	Intron			
4924	A	G	Intron			
4989	A	G	Intron			
5274	C	T	Intron			
5792	-	T	Intron			
5871	G	T	Intron			
6562	C	T	Intron			
9859	G	A	Intron			
9875	A	G	Intron			
10279	C	T	Intron			
10474	-	G	Intron			
10566	C	T	Intron			
12038	G	A	Intron			
12159	G	A	Intron			
12979	A	G	Intron			
13865	C	A	Intron			
13964	A	G	Intron			
14087	C	A	Intron			
14309	G	A	Intron			
16028	G	T	Intron			
16375	C	T	Intron			
16705	G	A	Intron			
19708	G	C	Intron			

21406	G	A	Intron
22401	T	C	Intron
22926	A	T	Intron
23007	G	A	Intron
23180	C	G	Intron
26490	A	G	Intron
26505	A	G	Intron
29336	T	C	Intron
29829	C	T	Intron
29830	G	T	Intron
29840	C	T	Intron
29944	C	T	Intron
30468	T	-	Intron
30471	T	-	Intron
30802	G	A	Intron
30894	G	A	Intron
30907	G	A	Intron
31447	C	A	Intron
31603	G	A	Intron
31685	A	G	Intron
31833	C	T	Intron
31970	C	T	Intron
32177	A	G	Intron
33018	T	C	Intron
33090	T	C	Intron
33993	C	T	Intron
34284	G	A	Intron
34314	G	A	Intron
35392	A	G	Intron
35599	T	A G	Intron
35997	C	A	Intron
36085	C	T	Intron
36270	C	T	Intron
36481	G	A	Intron
36619	G	A	Intron
37088	G	C	Intron
37204	G	A	Intron
37485	G	A	Intron
37624	-	A	Intron
37685	C	T	Intron
37769	C	T	Intron
38897	G	A	Intron
40155	A	G	Intron
40355	G	C	Intron
40486	T	C	Intron
40512	A	C	Intron
40622	-	A G	Intron
40654	T	C	Intron
40933	C	G	Intron
41171	T	C	Exon
41379	T	C	Beyond ORF(3')
41388	C	T	Beyond ORF(3')
41880	A	C	Beyond ORF(3')
42278	G	A	Beyond ORF(3')
42339	G	A	Beyond ORF(3')
42612	A	G	Beyond ORF(3')
42817	T	G	Beyond ORF(3')

418

L

L

Context:

DNA
Position
1573

TGGCTGTCCCCACACTGGAGATGCCCTACCTCCTGGTCTGGCCCACATGCAGTGGTGA
GCCCTCAGGGTCTTGTGACTTGGTCTATCCATGTGTCAGTCTGTAAAGGAGGACTTCT
GCCAGAACGTCCCCTCCAGGGCTGGAGCCATGACTCCCCTGTTACCCAACCTCAAGGT
GCCTGGCAGGAACCTCTATGATACCAGGCAGCCACAGAGGGAGGGATCAAAGTTGGAC
AGAGGCTGGTGTGAGAGACAGGATAGCCTAGACTGTGAACATGGGCAGTGGTTAGGGA
[A, T]
GTAGACATATGTGGTCAAACGTAAACAGAAAGCAAGGAAAAGGTACAAGCAACTCAGTA
CCTTAGGGAGAAGAGAATTAGGAGGGACACAGGGAGCTCAAACCTGGGAGTGTGTTG
TTCTTAAACTGGGCCATAAGTACATGGATGTGTGTTTATTATTCTTATACCTCAC
ATCTATTTACTCAGCAAATCTACAGAACCTCCTGTGTACCAGGCATTGTTCAAGTGC
TTAGAAATCTCTCTTAAGTAGATGTGATGGGTGTGAAATAATTGATGAAACAAA

1737

TACCCAACCTCAAGGTGCCCTGGCAGGAACCTCTATGATACCAGGCAGCCACAGAGGGAG
GGATCAAAGTGGGACAGAGGCTGGTGTGAGAGACAGGATAGCCTAGACTGTGAACAT
GGGCAGTGGTAGGGATGTAGACATATGTGGTCAAACGTAAACAGAAAGCAAGGAAAAGG
TACAAGCAACTCAGTTACCTTGTAGGGAGAAGAGAATTAGGAGGGACACAGGGAGCTTC
AAAAGGGAGTGTGTTGTTCTTAAACTGGGCCATAAGTACATGGATGTGTGTTTATTA
[T, G]
TCTTATATCTACACATCTTTACTCAGCAAATCTTACAGAACCTCCTGTGTACCAAGG
CATTTCAAGTGTGTTAGAAATCTCTCTTAAGTAGATGTGATGGGTGTGAAATAAT
TCATGATGAAACCAAGGGACACAGTAGGGCACTCATGTGAAAGAAGGGAGGCTAAG
GCATAGCATCAGAGGCCAAAATATCAGCTCCAACACCAGAGGATGCAATTCTTGTG
ATTAAACACTAAATTTCACTGCCAATTCAATTGCTCAGCTGAATAATCGGTTGCAGG

2498

AGGGCGGGAGGGAGCCTCTCCACCCCTCCCCTGAAACCTGGCAATCAGAACCCAGCCCCTG
ATGGAAGCCTGAGCTCTGGGGCCTCTGCCTCCCCCTCTTGTCAGCGTTTGTGTAAC
TGCCTCTGACCTCGGGAGAACCTCCAAGAGCTAGCCAGGGCTGGAGGCCTCAGCCGA
GCCTACCGGGTAAGACCAAGAAGGGCACCATGCTGTGCTCTCCCCTACGTTCACTCTA
ACACACAGCCCAGAGCCCCTAGAGGAGGCACACAGGGAGGAAAGCTGGTCAGGGATTG
[T, C]
GGGGAGACGGGGAGCAGCTGGTGCCTCTCTGTCTACGTGACTGTGGTGTCTCAGG
TGCCCTGGTGGAAATCATCCAGTAGGATCCAGGTGGAAAAGCCCTCATGCCAGCTAC
CGTTGAGGGCTTAACCCCAACTCCTGGCCGTAGCCCTGGATGCCCTAGGACCCACCTT
TCCCTCCCCACTCCACTCAAAGGCAGGTGCCAGCCTGGAGGTTCTCCAGGTT
TTTATCCCTTTGGACTTCTGCCTAGCCCTCAGAGAGAGTAGTCACTTACAATCAA

3097

AAAACAAAAAAGGTGACCCAACCTGTTCCAAATTCTCTGAAAGGGACTTGCCTCAGGT
GATTGTGTTCTCAAGGGAAAGGCTGAGTCGGCCCTCCATCCAGGGAGATGGACTGCC
ACCACCCCTACTCTTGCCTCACTGGTCCCTGGCCACCCAGGGCTGGCTGAAGACCC
TGTGCATGTGCCCCAGACGAGCATGCGCGAGCTGGTGGGTCAAGCCCTGAGACACGC
AGGAGGACTCTGTGCTGAGCAGCTCCAGAACATCCAGCTCCGTGCTGG
[C, T]
GGTATGTCCCCCTGCCCTGCCATGGTACTCTTGTAGAGGAAGAAATGCAAGGCAGAATT
GCCAGTTGCTTCCACGAGCATGTGCATAAAATGGAAAGACACAGCTCTCCAGACGCTG

3116

ACCTGTTCAAATTCTCTGAAAGGGACTTGCCTCAGGTGATTGTGTTCTCAAGGG
AAGGCTGAGTCGGCCCTCCATCCAGGGAGATGGACTGCCACCCACTCTTGCT
CACTGGTCTGGGCCACCCAGGGCTGGCTGAAGACCCCTGTGCATGTGCCCCAGAG
CAGCATGCGCGAGCTGGTGGGTCAAGCCGCTATGACACGAGGAGGACTCTGTGGT
GCTGCAGCCCTTCTCCAGAACATCCAGCTCCGTGCTGGTGTATGTCCCCCTGCC
[A, G]
CCCATGGTACTCTTTAGAGGAAGAAATGCAAGGCAGAATTGCCAGTTGCTTCCACGAGC
ATGTGCATAAAATGGAAAGACACAGCTCTCCAGACGCTG

4823

AACAAACTACTACACCCGTGTCTCTCTTCTCCCTGATCAGCTTGAACCACTTGGAG
AAAAACAGAGACCCCTGGACCTGAGAGCAGAGATGCCCATCACCTGTCCCACCTCAGGTAGT
AGGGAGGACCTGCCTGGCTCTCCACAAACCAAGGGCACACAGCTGCCCTACCCACT
TCGTCTCCACCAAGCTCCTCAGTACCCATCTGCCCTTACTGAGGCCTGAGGAGAT
TTGGAGGATGGAGGGAGTCCATGAGGATGGACAGGGAGGTGAGAGGGAGACAAGAGT

	[G, A]
	CACGTGTCATTGGAACAGGAGATGCAGCAGGGAGAGGAGGCCCTGGCCCCAAGCAGAGGG AGAGGATCCCGTGGAAAGTGGGCTCTGAGAGAGGAATCAGGATGCCAGGAAATG GCAGGAGGGCTCTCTTAGCAGTGGTGGGGCAGATGAAAAAATCTGACTGCAGGTT AGAGGGCCCAGGCAGGAGCAGGCAGGCTTAAGAGCTGTGGTGGAGAGAGGAGAGCCTG GATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGCAGCAAAGGGCAGAGCCCAG
4924	CCTGTCCTCACTCAGGTAGTAGGGGAGGACCTGCCTGGCTCCTCTCCACAAACCAGGGCAC ACAGCTCGCCCTACCCACTCGTCCTCACCACAGCTCCTCAGTACCCATCTGCC TTACTGAGGCTGAGAGATTGGAGGATGGAGGGAGTCCATGAGGATGGACAGGGAGG TGAGAGGGAGACAAGAGTCAGCTGCATTGGAACAGGAGATGCAGCAGGGAGAGGAG GCCTGGGCCAGCAGAGGGAGAGGATCCCGTGGAGAAAAGTGGCTCTGAGAGAGGAA [A, G] TCAGGATGCCAGGAAATGGCAGGAGGGCTCTTAGCAGTGGTGGGGCAGATGA AAAAATCTGACTGCAGGTTAGAGGGCCAGGCAGGAGCCAGGCTTAAGAGCTGTGG TTGGAGAGAGGAGACTGATTAGGGAGATTCACAAGGAAAGGATCACAGAGGACAGC AGCAAAGGGCAGAGCCAGAGCTGTATGGAGGAGGACGAGGGTGGGCTACCAGGACAC GGCAGCTCAGCTCTTTAAGGAGGAATCGTAAGTGGTGTAAAGCTTGACTTCAGG GATGCCAGGAAATGGCAGGAGGGCTCTTAGCAGTGGTGGGGCAGATGAAAAA [A, G] TCTGACTGCAGGTTAGAGGGCCAGGCAGGAGCCAGGCTTAAGAGCTGTGGTGG GAGAGGAGAGGCTGGATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGCAGCAA AGGGCAGAGGCCAGAGCTGTATGGAGGAGGGACGAGGGTGGGCTACCAGGACACGGCAG CTCAGGCTCTTTAAGGAGGAATCGTAAGTGGTGTAAAGCTTGACTTCAGG GGTGGGGCAGGTTCTCATGTCTCAGCTCCTGTTCTAGGCCGCTTATGGCTTT
4989	TCGCCCTACCCACTTCGTCCTCCACCACAGCTCCTCAGTACCCATCTGCC GAGGCCTGAGAGATTGGAGGATGGAGGGAGTCCATGAGGATGGACAGGGAGGTGAGA GGGGAGACAAGAGTCAGCTGCATTGGAACAGGAGATGCAGCAGGGAGAGGAGGCC GGCCCCAGCAGAGGGAGAGGATCCCGTGGAGAAAAGTGGCTCTGAGAGAGGAAATCAG GATGCCAGGAAATGGCAGGAGGGCTCTTAGCAGTGGTGGGGCAGATGAAAAA [A, G] TCTGACTGCAGGTTAGAGGGCCAGGCAGGAGCCAGGCTTAAGAGCTGTGGTGG GAGAGGAGAGGCTGGATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGCAGCAA AGGGCAGAGGCCAGAGCTGTATGGAGGAGGGACGAGGGTGGGCTACCAGGACACGGCAG CTCAGGCTCTTTAAGGAGGAATCGTAAGTGGTGTAAAGCTTGACTTCAGG GGTGGGGCAGGTTCTCATGTCTCAGCTCCTGTTCTAGGCCGCTTATGGCTTT
5274	GGGGCAGATGAAAAAATCTGACTGCAGGTTAGAGGGCCAGGCAGGAGCCAGGCC AAGAGCTGTGGTGGAGAGAGGAGGCCAGGCTGGATTAGGGAGATTCCACAAGGAAAGGATCA CAGAGGACAGCAGCAAAGGGCAGAGGCCAGAGCTGTATGGAGGAGGGACGAGGGTGGG TACCAAGGACACGGCAGCTCCAGGCTCTTTAAGGAGGAATCGTAAGTGGTGTAAAGC TTGACTTCAGGCTGGGTGGGGCAGGTTCTCATGTCTCAGCTCCTGTTCTAGGCC [C, T] GGTCTTATGGCTTTAACCAAATAAGGCCAAGGCCAGAAACCCCTCAGCAGCAATAAAA GCAGAAGGCCTGACCCAATCTGGGAGGCTGGGTTCCCTCTAGGTCGGCACACCACCC TCTCCCAACCCCTCCCTGCTGGGAATGGACCTGCAGCTCCCCATGTGCTGCTGGGAATC CTGAGAGAGTGGGCACCCCTGTTCACATGCCGCTCCCTGCTGCTGCCCTACCC AGTCTGGGCTCAGGCTCAGTCTGTGCCATCAGCCCCATCAGGAGAGCAAGAATGGC
5792	CTGCTGCTGCCCTACCCAGTCTGGCTCAGGCTCAGTCTGTGCCATCAGC CCATCAGGAGAGCAAGAATGGCAGGAAGAAGGGATGGGAAGTGAAGACAGTCGTAGCAG AGGGCTCAGTTGCTGGGTCTGTGCTGGAGCTAAGGAGATTGTCAGATTCTGCAACAGC TAGTGCAACACAGATGCCCTAGTCCAGGTGGTCAGGTGCTGCCAAAGGCCGGAGCAA AACCTTAGGCCCTACTGTGCCAGGTGAAACTCTTAACTGCTTCTAAGGATGCC [-, T] TGGGGGTTCTAGGGAGCAGCCAGGGACCGTGGATAGTGGGGCATTGGGACTCAGAA ATAGCCATATTGAGATATTCAATATTTCACCAACCCCTAGCCATACTGAATATCAGC CATGGAGGGCCCTTCAACAACTGTCCACTCCCCCTCCATTACATAACAAAAGCAGCCATC ATTGCTCTTCTTCACAAACAGTGTATTGAGTACTGAGTTGGAGCCTAAGCAGGGT CAGGGAGAGGCCCTGTCACCCCTGGCTCGAGGCAACCCTCAGGCTTACCCAGATC
5871	TGGCAGGAAGAAGGGATGGAAAGTGAAGACAGTCAGCAGAGGGCTCAGTTGCTGGGTC TTGTGCTTGGAGCTAAGGAGATTGTCAGATTCTGCAACAGCTAGTGCACACAGATGCC CTAGTCCAGGTGGTCAGGTGCTGGCCAAAGGCCCTGGAGCAAACCTAGAGGCCCTACT GTGCCAGGTGAAACTCTTAACTGCTTCTAAGGATGCCCTGGGGTCTAGGGGAGC AGCCAGGGACCGTGGATAGTGGGGCATTGGGACTCAGAAATGCCATATTGAGATA [G, T] TTCAATATTTCACCAACCCCTAGCCATACTGAATATGCCATGGAGGGCCCTTCAA ACTGTCCACTCCCCCTCCATTACATAACAAAAGCAGCCATCATTGCTCTTCTTCAC

FIGURE 3, page 18 of 33

D S D C D S D P D S D C D S D

		AAACGTGTATTGAGTACTGAGTTGGAGCCTAACGCACGGTCAGGGAGAGCCCTGTCA CTGGGCTCGAGGCAACCACCTCCAGGCTTACCCAGATCAGGCAGAGACCCCCAAAAG GAGGCTGCTCACCCAGCAGCATCTAAGCTGAGTGGCTCAGTGCCTCCCTCTAGACA
6562		CTAAGCAACATTGGAGCCCATTCTGAAAGGGTCCATCTGTTGGCCAGCCAACTTCACTG TGTCTGAGCATTCTGCATTCTCAGTCCCCTCTGCTCCCTCCATGTGCCTGGAGTGA TATAAAAGTCCACCAGCATTCTCAGTGTGAGCTGACAGGGGCCAGGCAGCACCTATTTTG TCCTAGATGTGTCTAACATAGAGGCAACAGGCAAGGCAAGACGCAGTGGGGGGCGGG AGGAGGAGGCCGAGATGGCTGTGAGCATGAGCTTCAGCCTCCCTCTCCATC [C, T] GCAGTCTAACTGCTCATACTGTTCTGTGTGCCAGGTAGGGTGACTTAACAGCACGCCATGG ATTCTGTTGAGTTCAAGTTGGACAAATTCTTACAGACAACCTTTGACTAGCCTTC TGTGGACTGAGCCTATACTCTGCCTTAATGGGCTCTGCCCACTCCTTCTAACCCCA GGCAGCCTGGCTGAACACCTGGCCTTTCTAGGTTCTTGTGACCTCTCTGAA GCCCTGTCAAAAGTCACCACCTCCCCCTGAAATTCACTCCTCCTGGGTTGTGGACA
9859		CAATAAAACCCAGCTAAAACAAGCCCAATAAAACCAATAAAACCCATTAGACAGGAACA TAGGAGTTGGAAAAAAAAGAAAAGAAGGGGAGGGGGAGAAAGCCCTGAGGCACCCCGGC TGCCGTCTGCCACAACCCCTGGCTGTAATTGTTCTGCCATGCCCTAGCTGCAACAC ATTCTAGTGTCTCCTGACCTCTAGCCCTAGCTGCTCCCTTCCCAACCTGTAG ATCTTGTGATCAAATAGATTCAATGAAACACATTGTCCAGTTG [G, A] CTCGCAGCACTCCAAAAGGTCAAGTTGTCCTCCCTCAGTGCCTCCATTCTGGTCA CGGTAGGACTGACTCCAGGCCCTGGACCCCTAAGCTGAGTCTGGCTCCTTGACGTGCAG GGAGAATGCCACTGAGTCTGTCTGAGGACCCCTACCTCTCAAATCTGCCCTAGTTC CTCAGCAGGTACTACACTGACTGCCATGCCATTCTCTGATGCTTCACTGCCCTAGCTTC TCAAGTCTGTCTCCCCACCTGAGCCAATTGTGAGTTCTCTCTC [A, G] CAATAAAACCCAGCTAAAACAAGCCCAATAAAACCAATAAAACCCATTAGACAGGAACA TAGGAGTTGGAAAAAAAAGAAAAGAAGGGGAGGGGGAGAAAGCCCTGAGGCACCCCGGC TGCCGTCTGCCACAACCCCTGGCTGTAATTGTTCTGCCATGCCCTAGCTGCAACAC ATTCTAGTGTCTCCTGACCTCTAGCCCTAGCTGCTCCCTTCCCAACCTGTAG ATCTTGTGATCAAATAGATTCAATGAAACACATTGTCCAGTTGCACTCGCAGCACTCCA CTGAGGTCAAGTTGTCCTCCCTCAGTGCCTCCATTCTGGTCACGGTAGGACTGACTCC AGCCCTGGACCTAAGCTGAGTCTGGCTCCTTGACGTGCAGGAGAATGCCACTGAG TCTGTCTCTGAGGACCCCTACCTCTCAAATCTGCCCTAGTTCCTCAGACTGCTCAGGTACTACA CTGACTGCCATGCCATTCTCTGATGCTTCACTGCCCTAGCTCTCAAGTCTGTCTCCCC ACCTGAGCCAATTGTGAGTTCTCTCTCCTCTCATCCTGGCACTAGAAATGCTC 10279
9875		CAATAAAACCCAGCTAAAACAAGCCCAATAAAACCAATAAAACCCATTAGACAGGAACA TAGGAGTTGGAAAAAAAAGAAAAGAAGGGGAGGGGGAGAAAGCCCTGAGGCACCCCGGC TGCCGTCTGCCACAACCCCTGGCTGTAATTGTTCTGCCATGCCCTAGCTGCAACAC ATTCTAGTGTCTCCTGACCTCTAGCCCTAGCTGCTCCCTTCCCAACCTGTAG ATCTTGTGATCAAATAGATTCAATGAAACACATTGTCCAGTTGCACTCGCAGCACTCCA [A, G] AAAGGTCAAGTTGTCCTCCCTCAGTGCCTCCATTCTGGTCACGGTAGGACTGACTCC AGCCCTGGACCTAAGCTGAGTCTGGCTCCTTGACGTGCAGGAGAATGCCACTGAG TCTGTCTCTGAGGACCCCTACCTCTCAAATCTGCCCTAGTTCCTCAGACTGCTCAGGTACTACA CTGACTGCCATGCCATTCTCTGATGCTTCACTGCCCTAGCTCTCAAGTCTGTCTCCCC ACCTGAGCCAATTGTGAGTTCTCTCTCCTCTCATCCTGGCACTAGAAATGCTC 10474
10474		GGGAGAATGCCACTGAGTCTGTCTGAGGACCCATTCTCCAAATCTGCCCTAGTT CCTCAGCAGGTACTACACTGACTGCCATTCTCTGATGCTTCACTGCCCTAGCT CTCAAGTCTGTCTCCCCACCTGAGCCAATTGTGAGTTCTCTCTCTCTCCTCTCATCCT GGCACCTAGAAATGCTCTAACGCTTGAGCTGCTAACCCAGCATGGTCACCTTGTTAT AGCAGTCTCCCAGATGCCCTTTGGTGAATGCTCAGGGAAATGCTTACTGTTAAC [C, T] GAGACAAGCCCAAGTAGCTACATGGACCTGCCACCATAGCCCTCTCCTGTCCTATGCTG TTGTAGAGGGTCCAGGGCTCACTCTCCACTTGCCCTGAGTACCTCTCCTTGAAAGGA TGTCAGGGGCTGGCGCACTGGCTAACGCTGTCACCTTGAGGACTTGGGAGGCTGAGGC GGCGGAGTCACCGGTCAAGGAGATCGAGACCATCTGGCTAACATGGTGAACCCCCCGTC TCTACTAAAAATAAAAAATAAGCCATTGTGGTGGCAGGTGCCTGTAGTCCCA TCAGGAGATCGAGACCATCTGGCTAACATGGTGAACCCCCCGTCCTACTAAAAATA AAAAATAAAAAATAAGCCATTGTGGTGGCAGGTGCCTGTAGTCCCACTGGGAGGCT GAGGAGGAAATGGCATGAACCCAGAAGGCAGAGCTTGAGCTGAGCCAGATCGCGCCA CTGCACTCCAGCCTGGCAACAGAGCAAGACTCCGTCTCAAAAAGCAAGCAAGAAAGAA AGGATATCGGTTACCTGTTAGACAGGAATGCTGAGACCAGGGAAAGGGGAGACTTGTG

FIGURE 3, page 19 of 33

	10566	CTTACTGTTAACCGAGACAAGCCCAAGTAGCTACATGGACCTGCCACCATAAGCCCTCT CCTGTCTTATGCTGTTGAGGGTCCAGGGCTCACTTCTCCCACCTGGCCCTGAGTACCT TCTCCTTGAAAGGATGTCAGGGCTGGCGCAGTGGCTCACGCTGTAAACCCAGCACTT TGGGAGGCTGAGGCAGGGGATCACCAAGGTCAAGGAGATCGAGACCACCTGGCTAACATG GTGAACCCCCCGTCTACTAAAAATACAAAAAATACAAAAAATAGCCATTGTTGGTGGCAGG [C, T] GCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAAAATGGCATGAACCCAGAAGGCA GAGCTTGAGCTGAGCCGAGATCGGCCACTGCACTCCAGCCTGGCAACAGAGCAAGACT CCGTCTCAAAAAGCAAGCAAGAAAGAAAGGATATCGTTACCTGTTCAGACAGGAATG CTGAGACCAGGGAAAGGGGAGACTTGTGCGGTGCTCAGGAAACCAGTATCTGAGCTGGG GGCTGAGAGCTGTGTTGGACTCTGCTCTCCAGTCGCTGCTGAGTCCCTCTCTTC
	12038	CTGGGAATACAGGCATGGGCCACTGTGCCAGCCTGGTTTTCTTCTTGTCCCATTAA TTCTCACATTTCAGACCATGGGCTACTACTCCACTGAGCACATTGTGAGAGTGCTC ACAGCCCTGGGCCGGTGTGTTCTGATCTCAGTCTTATCAACTTGATCTGCTTTG CTGTCATTATACATTTCTCATTAGCTTCTCCCATTCTCTTGTCTGCTTCCTTC TTCCTCTTAACTAACTCCTCACCTGCAACTGGGGGACTTGGATTCTGACTGGGCTT [G, A] TGTGAAAACGTATTGAAACAGATAGGTAAGTAGGAAATCAGGAGGGTGTGTTACAAGA AAAAAAAATGACTAAGATACAGGAACCAACCTAAAGAGGAAAGACATACAGTTCAAA GGAGGCAGAAAGAAAAACATTACAGATACTCAAATATATTGATAATCATAACACTTCTG GAAGATTAAAAAAATGCTGAAACATGAATCCCTGCTAGAGAAATTACAAAGCCAAGAAA ATAGATAGGTCTGAGGATTAGGGAGCTGTTCAAGTTGCTAGGAGGAACACAAAAGCACAGA
	12159	CAGCCCTGGGCCGGTGTGTTCTGATCTCAGTCTTATCAACTTGATCTGCTTTG TGTCAATTATACATTTCTCATTAGCTTCTCCCATTCTCTTGTCTGCTTCCTTC TCCTCTTAACTAACTCCTCACCTGCAACTGGGGGACTTGGATTCTGACTGGGCTT TGTGAAAACGTATTGAAACAGATAGGTAAGTAGGAAATCAGGAGGGTGTGTTACAAGA AAAAAAAATGACTAAGATACAGGAACCAACCTAAAGAGGAAAGACATACAGTTCAAA [G, A] GAGGCAGAAAGAAAAACATTACAGATACTCAAATATATTGATAATCATAACACTTCTG AAGATTAAAAAAATGCTGAAACATGAATCCCTGCTAGAGAAATTACAAAGCCAAGAAA TAGATAGGTCTGAGGATTAGGGAGCTGTTCAAGTTGCTAGGAGAACACAAAAGCACAGAC CCCAGACTACATGGGTATGAAACCTCTGCACGCCATTGTTGTCCATCCCTGCCAA GCTGTTATGTAACACCTCCGGGGAAATGAATTATGTTATACAGTTCTTCTAT
	12979	CAAACCATTTCCACCTGCCAGGGCTCGGGTGTGGTACAGGTTCAGAGTATTCACTGAA GCAGAAATGTACTTCTTACATACTGGGGATTGGAATGTACAGAAAAGGCTCCGGACCAC GAAGCCCCAGGATTGTCTAACATGTTCTCAAGTTGCTTACCTGACGTCAAGCCCCAAC AGAGGAAGTGTCTATGGATCGATTCTTGTACCTGGCAATCCTGGCTCACAGACGT GTTACTGCTTAGGCAGCTCAGCCTCTCAAGAGGGAGAGGAGCTGGTGTGATGTGGCGTT [A, G] ACTCTTGGAAAGTGGAGGCTGAGTGGGAGGAACTACAATTCTGGGATGGGACCCAAA AGGAAGTGGAGGCACGTTGTCATGTTCTGTGGCCCTAGGCCTTGTGTTGGTTCAAGT CAATCATCTAGTGTGAGGATTCAAGAGCCATGTTAACCTGGATTAAACCATGTC TGTGACTCTAGGACGGCACTGCAAAGACGGCTGGAGGACCCGACTATACCATGACT GGCAGTCAGGCCTGGTCGGGATCAGGTCTGTTGTCACCAAGGATGGGTTGACCCGAG
	13865	GGGCTGGGCAAGGGAGGAAATTCAAGGTGCACCATGCAGGCCAGGGCTTCTGAGGTGA GATTAAACTGAGACATGCATAATGAGGAGACACTTGCTATACAGGGAGCCAGGAACACA GTCCCAGGCAGAAGGACCATGGACCAACAGGCTCAGAAGTGGGACTGTGTTGGGTGTAT TTGGGGAAAGAGAAAGGCTAGAGTGGCTGGGGCATGAGAATGAGGTGGAGAGTGGGG GAAATGAGATCAGGAGTGCCAAGGAGGCCAGTCACACAAAGCCTGAATTACTGAGTAAA [C, A] CACTGGATTTCAGTGGAGAAAGATGGAAGGCATTGGCGTCTCAGGAGAGAGTGACAT GATCTGGTTCACTCTTCAAAGATCTCCCTGACTGCTATGTGAGAATGGGTTGGCCAT CAGCAGGAGTGAATTGGGAAAGACATTCTAAGCCAGCTGAAGAAACTAACCCATATGA AATCATTAAGAAACTATTGATGCTAAGCTCTGGGTGCAAGCAATACCAAGATTGCTGGCT GGGGTTATGCTGTCAGCCTCTGAATTCTCAGGCTCACGTTAGCCAGTGGAG
	13964	ATACAGGGAGGCAGGAACACAGTCCCAGGCAGAAGGACCACAGGCTCAGAA GTGGGACTGTGTTGGGTGATTGGGGAAAGAGAAAGGTCAGAGTGGCTGGGGCATG

D S C D C G C T - D S C D C G C T

	AGAATGAGGTGGAGAGTGGGGAAATGAGATCAGGAGTGCAAGGAGCCAGATCACACAA AGCTGAATTACTGAGTAAAACCCTGAGTTCAAGTGGAGAAAGATGGGAAGGCATTGG CGGTCTCAGGAGAGAGTGACATGATCTGGTCACGTCTTCAAAGATCTCCCTGACTGCT [A, G] TGTGAGAATGGGTTGCCATCAGCAGGAGTGATTGGGAAAGACATTTATAAGCCAGC TGAAGAAACTAACCCATATGAAATCATTAAGAACTATTGGATGCTAAGCTCTGGGTGCA AGCAATACCAAGATTGCTGGCTGGGGTTATGCTGTGTCAGCCTCTGAAATTCTCAG GTCACGTTAGCCCAGTGGAGGCTTGTCTCATTGAACCAGTGACCAAATTCCCTGAGAA TTGAAACGTCAGCTGCATCTGTGAATCAGGCATTCTTCATTATTACATTACATTG
14087	ATGAGGTGGAGAGTGGGGAAATGAGATCAGGAGTGCAAGGAGCCAGATCACACAAAGC CTGAATTACTGAGTAAAACCCTGAGTTCAAGTGGAGAAAGATGGGAAGGCATTGGCGG TCTCAGGAGAGTGACATGATCTGGTCACGTCTTCAAAGATCTCCCTGACTGCTATG TGTAGAATGGGTTGCCATCAGCAGGAGTGATTGGGAAAGACATTTATAAGCCAGCTG AAGAAACTAACCCATATGAAATCATTAAGAACTATTGGATGCTAAGCTCTGGGTGCAAG [C, A] AATACCCAGATTGCTGGCTGGGGTTATGCTGTGTCAGCCTCTGAAATTCTCAGGCT CACGTTAGCCCAGTGGAGGCTTGTCTCATTGAACCACTGACCAAATTCCCTGAGAATTG AAACGTCAGCTGCATTTGTGAATCAGCATTCTCATTTCATTTACCTATTGGAT GCCTATGTAGAGTGGGACTGCACTAAGTGCTCGTAGACAGTGGTGAGCCGAATGGGTC TGGATCTGCCCTCTGGGTTCTCAGTCTCATGCATCTTGCTTGTGCTGGAAGAGCT
14309	CATTTATAAGCCAGCTGAAGAAACTAACCCATATGAAATCATTAAGAACTATTGGATGC TAAGCTCTGGGGTGCAGCAAGCAATACCAAGATTGCTGCTGCGGGTTATGCTGTGTCAGCCT CTCTGAATTCTCAGGCTCACGTTAGCCAGTGGAGGCTTGTCTCATTGAACCACTG CCAAATTCCCTGAGAATTGAAACGTCAGCTGCATCTGTGAATCAGGCAATTCTCATT ATTCAATTACCTATTGGATGCATATGTAAGTGAGTGGGACTGCACTAAGTGCTCGTAGACA [G, A] TGGTGAGCCGAATGGGCTGGATCTGCCCTTGGTTCTTCAGTCTCATGCATCTTGCT TTTGTGCTGGAAGAGCTAAAATCCAGAGCTAGAAGGGCTGTGTTGTTAACAGC TTTCTACTCAAAGTAACCACAGAAACAAATTCTGTATCTGAGGTAACGTGAATGAGCC TAGAGGACATTAGTTAAGTGAATAAGTCAGGCACAGAAAGACAAACTACATGTTCT CACCATATGCGGAAGCTTAAGAAGTTGACTTCACAGAAGTAGAGTATAAATAGTGGTTAT
16028	ATATGCAAAAGACATACAGGTATCCAGAAAAGACAGGCAGAAACCAGGAGCTTACAAAT TTTAAATATTGTGTTATTATTCTAAAAATATTAAATTATTGTCTAGGTTCTACCAT TATAATTAGTGTCACTGTTAGCTTAATTTATAAACACACATACCTGTAATCTCATGTTAG GCATCCAAATGCTGTGTTCCCTGGAGACCACCTGTGAGGACTTCATGGTTCTTC CCTGTTGGGCAAGCCACTGGCTCCATTCAAAGCATAGATATATGGGATAAGAAAGGT [G, T] GTGTGGGGTCACATGTGGAGACATGCACTATGGGTTGTGCATAGGGTAGCTAGACAC ACCCATTCTCCCCCTTAATTCCCTCCTAGCCACCTATAACTCACAGTTCTTCCCT CACATGATCCTGTATGGTACTCATTCTAGCCTCCATAAAAATCCCTAGTGGTCT TCTGGGCTGAAGCTTATCTCCCTGACAATGAGTGGGACTGAAATCTTCTCCCTG TTGATTTAGAAGTGGGCACTGACTTCCTGTGACAGAGTGGAAAGGCTTCCAATAGTGGT
16375	GGGTAGCTAGACACACCCATTCTCCCCCTTAATTCCCTCTAGCCACCTATAACTC ACAGTTCTTCCCTCACATGATCCTGTATGGTACTCATTCTAGCCTCCATAAAAATC CCTTAGCTGGTTCTCTGGGCTGAAGCTTATCTCCCTGACAATGAGTGGGACTG AATCTTCTCCTGTGATTAGAAGTGGGCACTGACTTCCTGTGACAGAGTGGAAAGG CTTCCAATAGTGTCCAACCTCTGGTACTGAAAACATCATCATCTCCTCAATTAAAGGG [C, T] CTTGGCGAATATCAGGTTGGGGAGACCCCTGCAACACATACCCCTGGAGCTTAAAGCAGGA CTTGCTAATTCCCTGCACTGAGACCTAGATCCTGCGGCCCTGCCGCCACAGCTGGGCTT CCATGTTGGAGGTGCACAGAGCTCCATTGGATGCTACTTCTGTCTCCTTATAGTCCC GTGGCAGTCCCTAGGCTCCCTGCCAGTGAGGCAAGGTAGAGTCAGGGATTGGGATCTA CCTGCCTGTGCTACATGACCCCTGCACTGGAATTCTGGACCACCCCAATGTCATCA
16705	CTGCAAACATACCCCTGGAGCTTAAAGCAGGACTTGCTAATTCCCTGCACTGCA GATCCTGCGGCCCTGCCGCCACAGCTGGCTCCATGTGGAGGTGCACAGAGCTCTCCATT GGATGCTACTTCTGTCTCTTATAGTCCCAGTGGCAGTCCCTAGGCTCCCTGCCAG TGAGGCAGGTAGAGTCAGGGATTGGGATCTACCTGCTGTGCTACATGACCCCTGCA GAACATTCTGGACCACCCCAATGTCATCAGGCTTCTGAGGGTGGATGATAGCCATG

TOE90204 - 00000000

		[G, A] AACCCATTCCCAGTCCTGGTTGGTCTGAATGAATGGGAGGGCAAAACTGCTAAA GCCTTAAGCTAAAATAAGTACAATGGGAGCAGTGGACAGAGTTAGACTCTGGTA AAATGTGTACTTTAAGAGGTAGATACCCCCAGCCCCACAACCACCTCTGCTTGTCTC CCCTAGTCCACCAAGCTCCGACCAAGCAGACATCAAAGTGGTGGCCCTGGGTGACTCTC TGACTGTGAGTAGTGAGCCATGAACCAGGATGGGAGCTAGAGTCCAGGCCAGGCCCTGC
19708		CCTGCTGGAGGAGGGAGAGGGTTATCTGCAAGAAGGGAGTCAGCCAGGCCCTGAAA AGCCCCAGACTCTGTGTCACCCATGTCCCCACCCCTGCATGCTCATCTCAGTTACTG TGAGGGTCTCGAGGCTCACCTGTGCTTCTCCTCCTCCTCTAAAGACATT CTGAAGAAGTTCAACCCCTACCTCCTGGCTCTACCAGCACCTGGGAGGGACAGCA GGACTAAATGTGGCAGCGGAAGGGGCCAGAGCTAGGTGAGTAGATGCCGTACAGGAGGGC [G, C] AG
21406		AAGAATGAATCTCCATCTCCAAAAATAATAATAATTAAAAATAATAAGATAACAAAGG AATCAAAAGATGAACCTCCCTGGCCACGAAGAGCTTGACTCTAGGTAAGGAGGCTAAACA AATGGGAAATAACTTTCTGAAAAAGACAATGCTGGGTATGGCAACAATGCGAGTGCTTGC ATGGAGTACAATTAAAGAGAACAGAACAGACAGTATGAACGTGACTGTCTAAAGACAG ATGCAGACCCAGAAGGGACCCCTGAAATCATCCAGTCAACCTCTCCTTTAAAGATGG [G, A] AAAAGTCATCCTAGCAAGATTCAAGCAACTTGTACAAGCTAACAGCAAGTTGGTAGCAG AGCTGAAAGTAGAACCACTGGTCCCTGGGTAAAAAGGAAATGCAAGATGTGGATCA GGGAGCCCAGAGAGGAGGCTCAAGGGAAAGTAGGACTTGGCTGGGCTGAAGGATGGGA AGAAGATGGCTAGGAAGAGGGAGAACAGCGGCATTGTAACTTCCCCTCCTACCCACGAG GGCTTATTGCCATGGATTCTTAGTCACACCTTGAACCTGTTAAAGGTTAAAGGCAC
22401		AGGCCAAGGTAGGCAGATTGCCAGGGAGTTGGAGACCAGTCTGAGCAACTTGGTAAACC CTGTCCTACAAAAAATACAAAAAATTAGCCGGCATGGTGCATGGACCTGTAGTCCC GCTACTCGGAAGGCTGAGATGGAGGATCACTTGAGCTCAGAGGGTTGAGGCTGCAGTG AGCCGTGATCACACCACTACACTCCAGCTGGCAACAGAGCAAGACCTGTCTAAAAA GGATACAATTAAACATTGTACCTGTGAAATCATCACCACAATCAAGATGAAAATGTGTT [T, C] ATCACCCACAGGGAGTTCTCAGGCCCTGGTAATCTCTCCCTCTGCTCCTTCTGTC CCTACCTCACACCCAGGCACCACTAACCTTCTTCCATCAAATAGATTAGTTGCAT TTTAAAAAATTATATAATGGGATCAAAGAGTATATACTTTTATCTGACTTATTAG CAAATGATTTCGATGCATCCATGTTATTGGTATACCAATAGTTCGTCCCTTTAT GGCTGAGTGTAGTGTCCGGATTATCGCTCATCCAGAACACCAAATGGTATTG
22926		GTTGTCCTTTATGGCTGAGTGTAGTGTGTTCCGGCATCATCGCTCATCCAGA ACACCAATGGTATTGTTTATTTATGGCAGACATCAGGGATGAAGGGAGAACTAATCC TGTCCATCTGGTTATGGAGAGGGAGAAAAAAAGTAGGGAGATGGGAATGGTGC GGAATCTAAGTAACCACAGAAAAGAAAACAAAGGATTAAGGAGCAGAGCAGGGC TTAGAAGTAAAGGTTAAAGGAGTCATTAAGCCTGGAAAGGAGAAAAGTAGGGATAATTG [A, T] GAGCTGTGACTTTCTCAAATATAACAAAGTTATTTTAAACAGGCAACTGAAGAAGA AATGAACAGGCTTGGCTTACGAAGAAAGAGCTTGAGGAAGTATAAGGAAAGTCCCTGAG GGGAGGCTTGGAGGGATCCAACCCGAGTGGCCGATGAGACTATTGGGTGGCAGGGGCTA GATCAATGTGGCTCCAGGGTCCAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTCTT GAGAATGGAATGACCTTGTACTGGTAAACATCATTCTTGTAAACACCTCTTCTAG
23007		TTATGGCAGACATCAGGGGATGAAGGGAGAACTAATCTGTCATCCTGGTTATTGGA GGGGAGAAAAAAAGTGGAGGAGTGGGAATGGTGCAGGAAATCTAAGTAACCACAGA AAAGAAAAACAAAGATTAAGGAGCAGAGGAGCAGGGCTTAGAAGTAAAGGTTAAAGGA GTCATTAAGCTGGAAGGAGAAAAGTGGGATAATTGTGAGCTGTGACTTTCTCAA TATAACAAAGGTATTTTAAACAGGCAACTGAAGAAGAAATGAACAGGCTTGGCTTAC [G, A] AAGAAAGAGCTTGGAGGAGTATAAGGAAAGTCCCTGGGGAGGCTTGACGGGATCCCA ACCCGAGTGGCGATGAGACTATTGGGTGGCAGGGGCTAGATCAATGTGGCTCCAGGGTC CAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTCTTGTGAGAATGGAATGACCTTGT CTGGTAAACATCATTCTTGTAAACACCTCTCTTCTAGGGCAAAATCCCATGTCGTGAG TCCTCGCTCTGAGCCGGACTAACGCCCTCTCTACCCCCCACCTAGGGACATGCCA

23180	TAAAGGAGTCATTAAGCCTGGAAAGGAGAAAATGAGGGATAATTGTGAGCTGTGACTTT TCTCAAATATACAAAAGGTATTTTAAACAGGCAACTGAAGAAGAAATGAACAGGCTT GGCTTACGAAGAAAGAGCTTGAGGAAGTATAAGGGAAAGTCCCTGAGGGGAGGCTTGACG GGATCCCACCCGAGTGGCGATGAGACTATTGGTGGCAGGGCTAGATCAATGTGGCT CCAGGGTCCAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTCTTGAGAATGGAATGA [C, G] CTTTGTACTGGTAACATCATTCTTCTGAAACACCTCTCTTCTAGGCCAAATCCCCTG TCGTGAGTCCTCGCTCTGAGCCGGACTAACGCCCTCTCTCACCCCCACCTAGGG CATGCCAGCCAGGCCTGGACCTGGTAGAGCGAATGAAAAACAGCCCGTAGTACAGG CCCCCAGGCCACCCCTGAAAGGTGCCCATCTCTGCTGGCTGGGAGGGACAGCCCCAT AAGGGTCCCTCTCACACAGCACTTCTGCTTGGCTAGCCAAAAGATCCTCGGAGAAG
26490	CTCCCTCTGCTATAAAGCAAAGCCCTGAGATTCAAGCCTGCAAGGACTTACTGAGCACCTA CTATGTACCTTGTTCATCACCCAGGATGCTGTGGACACACCTCTAAATCAGCCTCCTA CTGGGGAGATGGTCAGAGGAAGAGAACCTTACACTGAGTCACAGGGATAGAAGTTAGG GGAACACAGGAGAGCAAACATTCAGGCAGTGGGACCAGCATGGACCAAAGCCCAAAGG AAAAAGGAAGTGTGGCCACCCAGGGCATGGCAAGGGCTGGAGAAGGCTGAGGTAGATG [A, G] CGGATGGGACTGCCAAGAGCCAAGGCCAAAAAGTGGCAGGACCCAGCACTGGCAGACTCC ACTGTTGGGCTGAGATTATGTAGAGCAGGGTGGGGTTGGGATTGTTCATGGTGTCTAG TAGGGGACAAGGGATGATTCTTACAGAGACTCAGCAGCAACAAAGAAACTGGGCTTCTCAG TTGACCAGGACCACCGAAGCCCTGTCACCCACTCAGTCATTAGCCCAGGCCCCAGA GCCCTCTATGCTTGCCTACTCAGAGCGGGCACCAGGGCTAAAGAGAGTACCCCT [A, G]
26505	AGCAAAGCCCTGAGATTCAAGCCTGCAAGGACTTACTGAGCACCTACTATGTACCTTGT GCATCACCCAGGATGCTGTGGACACACCTCTAAATCAGCCTCTACTGGGAGATGGTTC AGAGGAAGAGAACCTTACACTGAGTCACAGGGATAGAAGTTAGGGAAACACAGGAGAGC AAAACATTTCAGGCAGTGGGACCAGCATGGACCAAAGCCCAAAGGAAAAGGAAGTGTGG CCACCCAGGGCATGGCAAGGGCTGGAGAAGGCTGAGGTAGATGACGGATGGGACTGCC [A, G] AGAGCCAAGGCCAAAAAGTGGCAGGACCCAGCACTGGCAGAGTCCACTGTGGGCTGAG ATTATGTAGAGCAGGGTGGGGTTGGGATTGTTCATGGTGTCTAGTAGGGACAAGGGAT GATTCCTTACAGAGACTCAGCAGCAACAAAGAACTGGGCTTCTCAGTTGACCAGGACAC CGAAGCCCTCTGTCACCCACTCAGTCATTAGCCCAGGCCCCAGAGCCCTCTATGCTCT TGCATTCTCTCAGAGCGGGCACCAGGGCTAAAGAGAGTACCCCTTTTCTTACAGGA
29336	AAATACTTACCCCTGCAAATTGAACACCAAGGCCAGGAAGGGAGTGAGAGACCCCAAAG TGGAAAGCTGAGAAAATCCCTCTCCAGCGGGTAGGCAGCAAGAGATTCCCAGAGTAGA CTCCTTGTGGTAGGGCCATTCCCCACCCAGAGCCATGTGTAATAATTACTACTCACTC CTCCCCCTCCCTCATTAAAACAAAAGGCTTAGGCCAGACAAATGGCTCACGTCTGGT TCCCAGCTACTCAGGAGGCTGAGATGGGAGGACAGCTTGAGGCCAGGAGTTGGAGGCTGC [T, C] GTTAGCTATGATGATGCCATTGTAECTCTGCCTAGACAACAGCGTGAGACCTATCTCAA AAAAAAAAAGAAAAAGAAAAAGGCTTAGGCCCTACTTAACTCTACCTCAAAT TCTCTTGCCTCTCTGCCCCCTCCATCTCCCCACCTCCACTCTGCTTATGTCT GCCTCTATTGTTCCCTCTCAGGCTCAGGTAGCATTCATTGCAAACAGCCCTCT CATTACAAGGCAAGTGTGCTTCCCTCTCAAGGAGCTTCCCTGCCTGAACCTCACCC
29829	CCCTCTCAGGCTCAGGTAGCATTCCATTCTGCAAACAGCCCTCCTTCATTACAAGGC AAAGTCTGCTTCCCTCTCTAAGGAGCTCCCTGCCACTAACCTCACCGCGGACATCTCC CCATATCACATTCACTCAGTCTGACTTGTAGGGCTAAAGGCCCAAAGGGTTCTCATGTT TCACATCTGCTCATTTTCCAGATGGATGATAAAACTCCTTGAAGATAAGTACATCTAG TCTGTTCTTACATTCCATGCTTGGTACTTAAATCCAGCCACCGTGACTCTCC [C, T] GCAAAGTTCATGGGCAATTGGGAGCTGGTGTGAGATGCTCCCCATCTGACCTGCAGCC CCATGTTCTAATTGACCTCTCGTGCAGTGAGAGGAGGGAGACTTGGCCTATGCAAT CTGGTCAGTGGCTCAGACCCAGCCTTCAGGCAGAGGCTTGGAGATGGGACTGGGAG CTGTGTAAGCTAGGGAGCTCTCCCACAGGAGCCCTGGGTTCAACTCATCTGATCCT GAGAACAGCATAGGGCTTGAAATGTCCGTGCCCATGAATGGTGGAGAATAAGTAT
29830	CCTCTCAGGCTCAGGTAGCATTCCATTCTGCAAACAGCCCTCCTTCATTACAAGGC AGTCTGCTTCCCTCTCTAAGGAGCTCCCTGCCACTAACCTCACCGCGGACATCTCC CATATCACATTCACTGCTGACTTGTAGGGCCCTAAAGGCCCAAAGGGTTCTCATGTTT

FIGURE 3, page 23 of 33

Sanger sequencing

	CACATCTTGGCTCATTTCAGATGGATGATAAAACTCCTTGAAGATAAGTACATCTAGT CTGTTCTTTACATTCCATGCTTGGTACTTAAATCCAGCCACCGTGGACTCTCCTCCC [G, T] CAAAGTTCATGGGCATTGGGAGCTGGTGTGAGATGCTCCCCATCTGACCTGCAGCCC CATGTTCTAATTGACCTCTCGTGCAGTGAGAGGAGGGAGGACTTGGCCTATGCAATC TGGCAGTGGCTCAGACCCAGCCTTCAGGCAGAGGCTTGGATGGACTGGGAGCTGGAGC TGTAGCTAGGGAGCTCTCCCACCAGGAGCCGCTGGTCAACTCATCTGATCCTG AGAACAGCATAGGGCTTGAATGTCCGTGCCCATGAATGGTGGAGAATAAAAGTATG
29840	TCAGGTAGCATTCCATTCTGCAAACGTACCCCTCCATTACAAGGCAAGTCTGCTTC CCTCCTCTAAGGAGCTTCCCTGCTGTAACCTACCCGCGGACATCTCCCCATATCACAT TCAGTCTGACTTGTAGGGCCCTAAAGGCCAAAGGGTCTCATGTTTACATCTGG CTCATTTCAGATGGATGATAACTCTTGAAGATAAGTACATCTAGTCTGTTCTT TACATTCCATGCTGGTACTTAATCAGCCACCGTGGACTCTCTCCGCAAAGTTCA [C, T] GGGCATTTGGGAGCTGGTGTGAGATGCTCCCCATCTGACCTGCAGCCCCATGTTCTAA TTGACCTCTCGTGCAGTGAGAGGAGGGAGGACTTGGCCTATGCAATCTGGCAGTGG CTCAGACCCAGCCTTCAGGCAGAGGCTTGGATGGACTGGGAGCTGTAGCTA GGGAGCTCTCCCACCAGGAGCCGCTGGTCAACTCATCTGATCCTGAGAACAGCA TAGGGCTTGAATGTCCGTGCCCATGAATGGTGGAGAATAAAAGTATGTTGCATCCC
29944	TCTCCCCATATCACATTCACTGTACTTGATGGGCCCTAAAAGCCCCAAAGGGTTCTCA TGTTCACATCTGGCTCATTTCAGATGGATGATAAAACTCCTGAAAGATAAGTACA TCTAGTCTGTTCTTTACATTCCATGCTTGGTACTTAAATCCAGCCACCGTGGACTCT CCTCCCGCAAAGTTCACTGGCATTGGAGCTGGTGTGAGATGCTCCCCATCTGACCT GCAGCCCCATGTTCAATTGACCTCTCGTGCAGTGAGAGGAGGGAGGACTTGGCCTA [C, T] GCAATCTGGTCAGTGGCTCAGACCCAGCCTTCAGGCAGAGGCTTGGATGGACTGG TGGAGCTGTGTAGCTAGGGAGCTTCTCCACCAGGAGCCGCTGGTCAACTCATCTG ATCCGTGAGAACAGCATAGGGCTTGAATGTCCTGCCCCATGAATGGTGGAGAATAAA AGTATGTTGCATCCACTAGACTAGCCCTAAAGTCACTGTCCTTAGGGTGAGTTGA CTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCTGTGAGCCTTGCCAG
30468	CCTTTAGGTGAGTTGACTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCCTGT CTGTGCAGCCTGCCAGGAGGGTTGAGCAGCTCTCTCTGTCCCCAGCATGGCATCTC CAGTTCTCCTACTGGCACCAATACACACAGCGTGAGGACTTGCAGGTTGGTGCAGCC TTCTTCCAAAACACACTCACCCCACTGAACGAGGTGAGCTGCAGGTATTTAGGGAGGC TCACGTATGGGGCCTTATCACAGACGATGGATGTATTCCTCTTAAGTGGCTTTTT [T, -] TTTTTTTAACCATCTCTCCAAGAGGATTCTGAGGGTGGCTTTCCACATTACCTC CTTTTGTGGGGCTGGGCTGTGATTGAACTCAGATGTACTTGAAGGAAATCAATAG TGACTAAGCTCCAGGCCTGGCCCTGATGTTTCTGGATTGGATAGAATGGAAAGCTTC CTAAAAATGTTACTCTTTCAACTCTTAGGATAGGGTGCTGAAAGAAAAGGGAGAGACT ATGGGTGGGTCATTCTGCTGTTAAAAAGAAAATCCGGCCGGGTGCAGTGGCTCA
30471	TTAGGGTGAGTTGACTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCCTGTCTG TGCAGCCTGCCAGGAGGGTTGAGCAGCTCTCTCTGTCCCCAGCATGGCATCTCAG TTCTCCTACTGGCACCAATACACACAGCGTGAGGACTTGCAGGTTGGTGCAGCCTT CTTCCAAAACACACTCACCCCACTGAACGAGGTGAGCTGCAGGTATTTAGGGAGGC CGTATGGGGCCTTATCACAGACGATGGATGTATTCCTCTTAAGTGGCTTTTT [T, -] TTTTTAAACCATCTCTCCAAGAGGATTCTGAGGGTGGCTTTCCACATTACCTC TTGTGGGGCTGGGCTGTGATTGAACTCAGATGTACTTGAAGGAAATCAATAGTG CTAAGCTCCAGGCCTGGCCCTGATGTTTCTGGATTGGATAGAATGGAAAGCTTC AAAATGTTACTCTTTCAACTCTTAGGATAGGGTGCTGAAAGAAAAGGGAGAGACTATG GGGGTCCAATTCTGCTGTTAAAAAGAAAATCCGGCCGGGTGCAGTGGCTCATGC
30802	TGAGGGTGGCTTTCCACATTACCTCTTTGTGGGGCTGGCTGTGATTGGAACCTC AGATGTAATTGAAAGGAAATCAATAGTGACTAAGCTCCAGGCCTGGCCCTGATGTTT CTGGATTGGGATAGAATGAAAGCTTCTAAAGTGTACTCTTCAACTCTTAGGATA GGGGTGCAGGAAAAGGGAGAGACTATGGGTGGTCCAATTCTGCTGTTAAAAG AAAATTCCGGCCGGGTGCAGTGGCTCATGCCTGTAATCTCAGCCTTGGGAAGCCAAGGC [G, A]

D E S C R I P T I O N

	GGAAGAAGGGACCTTATGAAATGAACATACAGTCTGGGGTCTTCAGGGACACCTGCCT GGTCTTCACTCTGCCCTCTGTGGCTGCCACCAGCAACTGAACGGTTCCGCACAGCA
31833	ACTGCGAGGGGATCCCTTGTAAATCAGATAGGTTGGCTAGATGAAAATACCAACTTCTAC CTCGTACTGTGTGACCTTGGCAAACGATCTCTGGCCACCTGTATCAACATCTATAAA ACAGTAAAACAAGACAGGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAG CACAATAGTTAGCACTCAGCAAATGTCAACCACCATCAGCCTTCCAAGCACTCCGGGCTCA ACTCATACCCAACTCATTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTTC [C, T] GAGGCTGATAACCTATTCCAGTCCTTCTGATGGGAAGAAGGGACCTTATGAAATGAACAT ACAGTCTGGGGTCTTCAGGGACACCTGCCTGGTCTCCACTCTGCCTTCTGTGGCTG GCCACCAGCAACTGAACGGTTCCGCACAGCACTTGACCTGTCACCCCAACAACTGGAT CCTCTTGACGGAGCAAATGAAATGCCCTCCCAACCCAATGGTTCTTTAATCCAGGCT CAGTGGGTAACACAATCCCCACCCCAACCTGTATGTTCCCTCTGTGCTATGACAAC
31970	GGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAGCACAATAGTTAGCACTC AGCAAATGTCAACCACCATCAGCCTTCCAAGCACTCCGGCTCAACTCATACCCAACTCAT TTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTCCACTCTGCCTTCTGTGGCTG TCAGGGACACCTGCCTGGTCTCCACTCTGCCTTCTGTGGCTGCCACCAGCAACTGAA [C, T] GGTTTCCGCACAGCACTTGACCTGTCACCCCAACAACTGGATCCTTGCACGGAGCAA ATGAAATGCCCTCCCAACCCAATGGTTCTTTAATCCAGGCTCAGTGGTAACACAATC CCCACCCCAACCTGTATGTTCCCTCCTTGCTCTATGACAACAAAGCTACATTCCA GCTCCTTTTATCACAGTTCAAGGCCGTAGTGTCTCTGCCAACCCAGCTGTGCAAACGT TCCCACCCCTGTCACTCCAGTATGTCAGCATCCCACACTCGGCTGACTCACAAATT
32177	CCTTATGAAATGAACATACAGTCTGGGGTCTTCAGGGACACCTGCCTGGTCTTCCAC TCTGCCTTCTGTGGCTGGCACCAGCAACTGAACGGTTCCGCACAGCACTTGACCTGTC ACCCCCAACAACTGGATCCTCTTGACGGAGCAAATGAAATGCCCTCCCAACCCAATGGT TTCTTTAATCCAGGCTCAGTGGTAACACAATCCCCACCCCAACCTGTATGTTCCCTCC TTTGCTCTATGACAACAAACAGCTACATTCCAGCTCCTTTATCACAGTTCAAGGCC [A, G] TAGTGTCTCTGCCAACACCAGCTGTGAAACGTTCCCACCCCTGTCACTCATCCAGTAT GTCCAGCATCCCACTCGGCTGACTCACAAATTGACTTTCTCCTAGCTATACCATCTCC TCCTCTCTAGCAACCTCTCTTTAAGAACAGCATGTAACACTGGTTATCCTGGCCTA GTTAATGGCAGACTCAGTTATGTCGACTTCAATTGTGAGGGGTTTCTCTGTGGAC ATCACGTACCTGCCCACTCCAAGAACTTCTATTGTAATTGACTCTTCAGCCAAGACTCCGGAT
33018	AAGGTATGCCCTCCTACCAAGGTGGCACTCCAAGTCTGTTAAATCTGGGACCCCTCCAGG AATCTCCTGGGGCTGGATAGCCATAGTGAACGGCTGAACTGAAAAAGAGTCCATTGGTT TCTTTCTGTGAATTAAACAATGTAGCTCTGCCAGGCACGGCTCATGCCTGTAATC CCAGCACTTGGGAGGCCAGGCAAGTGGATCGCTTGAGCCAGGAATTAGACACCAACC TGGCAACACAGGGAGATTCTGTCTACAAAATAATCAAATATTAGCCAGGTGTGG [T, C] GGTGCATGCCGTAGTCCAGCTGCTCAGAAGGCTGACGTGAGAAGATCACTGAGCATG GGAGGTCAAGGCTGCAATGAGCCGAGATGGCACCACCGCACTCCAGCCTGGCAATAGAG TGAGACCCCTATATCTCAAAAACAAATAGAAAAAAAAATATGTAGCTCTGGCCTCT CTTCTAAAGCAGTTCACTGAGCTCTCCATTCAACCCAGGTAAGAGGCCCTTATTTCATAA AGATAAGTGGAGGAGTTAGATATGAAAACAAAAGTAAACACCGCACTGGAGCTATTG
33090	CTGGATAGCCATAGTGAAGGCTGAAACATGAAAAAGAGTCCATTGGTTCTTCTTG AATTAACAATGTAGCTCTGCCAGGCACGGCTCATGCCTGTAATCCAGCACTTGG GAGGCCGAGGCAGGTGGATCGCTTGAGGCCAGGAATTAGACACCAACCTGGCAACACAG GGGAGATTCTGTCTCTACAAAATAATCAAATATTAGCCAGGTGTGGTGCATGCT GTAGTCCCAGCTGCTCAGAAGGCTGACGTGAGAAGATCACTTGAGCATGGAGGTCAAGG [T, C] TGCAGTGGCCAGGATGGCACCACCGCACTCCAGCCTGGCAATAGAGTGGAGACCCCTATA TCTCAAAAACAAATAGAAAAAAAAATATGTAGCTCTGGCCTCTCTCTAAAGCAG TTCAGTAGCTCTCCATTCAACCCAGGTAAGAGGCCCTTATTTCATAAAAGATAAGTGGGA GGAGTTTAGATATGAAAACAAAAGTAAACACCGCACTGGAGCTATTGTGAAACAAAAC AAGACTGTCCATGGTCCCCAGCCATTATTATCTCAGCCATACCCGAATTCAAATAA

33993	ATTACAGACCAGGGTTTCTAGTCTTCTGGTACCTGGCATGCCACCACCCCTCCCC ACTGCTCCCAACCTGATAAGCACATATATACCGGTGAATTCTCATGTCTCACAATTAGAGT CCTATGACATAGTGTCTGCAGGGCTTGGCTGATGTCCTCATAGTGTCTGCAGGGCTTGGC TGATGTTCCCAGGGTCCCTACTAGGAAGCAAAGCACCTAAACTATTTCATCTTATT TCATCTCCTGCCCTCCTCTCACGTCTTCAGAGACTTTGCAAAGGCAAAGCCAGAAG [C, T] TCCAGCAGCACCAGGGATATTCCTCTGCCTCTGCTTCTGTCTTCTGAA GAAGTTCTTCCCAGGCCTAGTCCTACTGCTGCCTACTCCCTCTGCAGA AATCTGCTCAGCCAGTGTGATCTCCCAGGGCTGGGTGACAGCTCAGCCTCC TAAGTACATCCCTGTCTCAGACTTAGAGCTCTAGAAATCGTACTCAGCTGGCT GCATATTAGAATCATTAGGGACATTGTGTATGTGTATGTATGTATATGTATGA
34284	AGCCAGAAGCTCCAGCAGCACAGGGATATTCCTCTGCCTCTTCTGTCTTCTGAA CTTATCTGAAGAAGTTCTTCCCAGGCCTAGTCCTACTGCTGCCTACTCCCT CTTCTGCAGAAATCTGCTCAGCCAGTGTGATCTCCCAGGGCTGGGTGACAGC TCCAGCCTCTAACGTACATCCCTGTCTCAGACTTAGAGCTCTAGAAATCGTACTCTC AGCTCTGGCTGCATATTAGAATCATTAGGGACATTGTGTATGTGTATGTGTAT [G, A] TATGTATGAATGTGTGTGTATGTGTGTATGTGTATGTGTATGTGTATGTGTATGTA TGTATGTATGTATGACAGAGCTCACTCTGGCCAGGGTGGAGAGCAATGGCACCATC TCAGTCACTGCAACCTCCGCTCTGGATTCAAGCATTCTGCCTCAGCTCCCAA GTAGCTGGGTTATAGGTGCATGCCACCATGACCAGCTAATTGTATTTAGTAGAG ACAAGGTTCGCCATGTGGCCAGGGCAGGGCTGGTCTTAACCTGACCTCAGGAGATC
34314	TTTCTCTCTTCTGCCTCTTCTGTCTTCTATCTGAAGAAGTTCTCTTCCCAGG CCTAGCCTCTACTGCTGCCTACTCCCTCTGCAGAAATCTGCTCTCAGCAGT TTGTATCTCCCAGGTGCTGGGTGACAGCTCAGCCTCTAACGTACATCCCTGCTTC AGACTTAGAGCTCTAGAATCGTACTCTCAGCTCTGGCTGCATATTAGAATCATTAGG GACATTGTGTATGTGTATGTGTATGTATGTATGTATGAATGTGTGTATGTGTGT [G, A] TGTATGTATGTATGTGTATGTGTATGTGTATGTGTATGTATGACAGAGCTCACTCTG TTGCCAGGGTGGAGAGCAATGGCACCATCTCAGTTCACTGCAACCTCCGCTCTGGAT TCAAGCAGATTCTCTGCCTCAGTCTCCAAAGTAGCTGGGTTATAGGTGCATGCCACCAT GACCAGCTAATTGTATTTAGTAGAGACAAGGTTGCCATGTGGCCAGGCCAGG CTGGCTTTAACCTGACCTCAGGAGATCCACCCACCTCGGCCTCCAAAGTGTGGGA
35392	GACGTCAAGCATGATGGGTAGAGACTCACTGGGGATGAATAGTCTGGAAAGAAGGTGGA AAGGGCTTTGAGGACTATAATAGTCTGTTGCCTGACTGGATGCTGGTATGTCATT ATCGAAACTTATCTGTTGCTCACTTATGATTGTTACTGTTCTATGTGTATGTTAGCTT CAATTAAAAGTTACTTGAGGCCGGGTACAGTGGCTCACACCTGTAATCCCAGCACTTTG GGAGGCCAGGGCAGGAGATCCCTGAGGTCAATACCAGCTAGCCAACATG [A, G] TGAAACCCCATCTACTAAAAAATACAAAATTAGCCAAGCGTGGGCCACGTGCCTATAA TTCCAGCTACTTGGGAGGCTGAGACAGGAAATCGTTGAAACCCAGGAGGCCAGGGTTG AGTGAGCCAAGATGCACTTGCACCTCAGCCTGGGTGACAAGAGTAAAACCTGTCTC AAATTAAAAAAAAAAAAAGTTACTTGAAAACAATATCAGTGCCTGACCGG GCTTATCCCCAGAGAGTCTGACTTAATTGGTCTGGAGTGCAGCTGGATTGCGTACTT GTTACTTGGAGGTTCAATACCAGCCTAGCCAACATGATGAAACCCATCTCTACTAAA AAATTAGCCAAGCGTGGGTGACGTGCCTATAATTCCAGCTACTTGGGAGGCTGAGACA GGAAACCGCTTGAACCCAGGAGGCCAGGGTTGAGTGCAGGCAAGATTGCACTATTG TCCAGCCTGGGTGACAAGAGTAAAACCTGTCTCAAATTAAAAAAAAAAAAAAA [T, A, G] TTTACTTGGAGGTTCAATACCAGCCTAGCCAACATGATGAAACCCATCTCTACTAAA TGGTCTGGAGTGCAGCTGGATTGGTACTTTGTGAAAGCTCTGAGATTATTTAATG GCAGGGTTATGAACCGCTGCCATTAGATCTGGTCCCCACAGAGAAATCAAGTAATCTG TAAAAAGAAAACCTGACCCAGTCAGTCACTCCCTGCTTCAAACCTCCAAAGCCTCCAC GAAGGAGGCAGGCCAGGCCAGGCCAGGCACACTAGGCCTCTGGGACTTGGCCTGGTT
35599	ACAGTGGCTCACACCTGTAATCCCAGCACTTGGGAGGCCAGGGCAGATCCCTGA GGTCAGGAGTTCAATACCAGCCTAGCCAACATGATGAAACCCATCTCTACTAAA AAATTAGCCAAGCGTGGGTGACGTGCCTATAATTCCAGCTACTTGGGAGGCTGAGACA GGAAACCGCTTGAACCCAGGAGGCCAGGGTTGAGTGCAGGCAAGATTGCACTATTG TCCAGCCTGGGTGACAAGAGTAAAACCTGTCTCAAATTAAAAAAAAAAAAAAA [T, A, G] TTTACTTGGAGGTTCAATACCAGCCTAGCCAACATGATGAAACCCATCTCTACTAAA TGGTCTGGAGTGCAGCTGGATTGGTACTTTGTGAAAGCTCTGAGATTATTTAATG GCAGGGTTATGAACCGCTGCCATTAGATCTGGTCCCCACAGAGAAATCAAGTAATCTG TAAAAAGAAAACCTGACCCAGTCAGTCACTCCCTGCTTCAAACCTCCAAAGCCTCCAC GAAGGAGGCAGGCCAGGCCAGGCCAGGCACACTAGGCCTCTGGGACTTGGCCTGGTT
35997	AGCTCTGAGATTATTTAATGTGCAGGGTTATGAACCGCTGCCTAGATCTGGCCCC ACAGAGAAATCAAGTAATCTGTATAAAAGAAAACCTGACCCAGTCAGTCACTCCCTGCTT AACTTCCAAAGCCTCCACCTCTGAAGGGAGGCCAGGCCAGGCACACTAGGCCTCTGG GAAGGAGGCAGGCCAGGCCAGGCCAGGCACACTAGGCCTCTGGGACTTGGCCTGGTT

S 3 2 7 3 3 3 4
= 0 2 0 6 0 4

	GGCCTCTGGACTTGGCTGGTCACCTGATTAACCTCTGGTACCCATTCCACCAGC GTCTGCCTCGCATGTTACAGTCTAGTGAUTCCAGCAGCGTCCTGCACCACTGTGGTGT [C, A] CACACCTCTGCTAATCTTGCTCTCCTCCTCTGGATTGCCCTCTCACCTCCTGC CCACTCCACCACTCAACTCAGGTGCCACCTCCTGCAGGAAGCTACCTCTGAATCTCAGG ACAGGCCAGTGGCCACCCAGGTCCATTACACCCCTGCCAGTCTGTCAATTGCTACGTG GTTGGTAGCCACAGTGCCTGGCTAGGAAAGACTGGTCTAGGAAAACAATTTCATTCC CTGTGGCCAGCTCCAAGCCTTCCCCGCCAACCTCTCCATTCAAGGTCTGTGAATTTA
36085	GAAAACCTGACCCAGTCACTCCCCCTGCTTCAAACTTCAAAGCCTCCCACCTCTGAAGG AGGCAGGCCAGGCCCCATAGCACAGCACACTAGGCCTTGGACTTGGCCTGGTCAACCT GATTAACCTCTGGCTACCATTTCAACAGCGTCGCTCGCATGTTACAGTCTAGTGA CTCCAGCAGCGCCTGCACACTGTGGTGTCCACACCTCTGCTAACTCTGCTCTCC CCTCTCTGGATTGCCCTCTCACCTCTGCCACACTCAACACTCAGGTGCCA [C, T] CTCCCTGCAGGAAGCTACCTCTGAATCTCCAGGACAGGCCAGTGGCCACCCAGGTCCATT ACACCCCTGCCAGTCCTGTCAATTGCTACGTGGTGTAGGCCACAGTGCCTGGCTAGG AAGACTGGTTCTAGGAAAACAATTTCATTCCCTGTGGCCAGCTCCAAGCCTCCCCGC CAAGCTTCTCATTCAAGGTCTCTGTGAATTAAATTAACTCCATCCATCAAACAAGTA TTTACTGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGAAGTCAGCAGTGAAA
36270	GCAGCGTCTGCACCACCTGTGGTGTCCACACCTCTGCTAACTCTTGCTCTCCTCCTC TCCCTGGATTGCCCTCTCACCTCTGCCACCTCCACCACTCAACTCAGGTGCCACCTCC TGCAAGGAAGCTACCTCTGAATCTCCAGGACAGGCCAGTGGCCACCCAGGTCCATTACAC CCTGCCAGTCCTGTCAATTGCTACGTGGTGTAGGCCACAGTGCCTGGCTTAGGAAAGA CTGGTTCTAGGAAAACAATTTCATTCCCTGTGGCCAGCTCCAAGCCTCCCCGCCAAG [C, T] TTCTCCATTCAAGGTCTCTGTGAATTAAATTAACTCCATCCATCAAACAAGTATTAC TGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGAAGTCAGCAGTGAAAAGATG ACTGCTACTCTCATGGGACATACAGGAGTAGTAGGAAAAGACAGATAATCAACAAGGTCA TTTCTGACCACATCTGTGGTTAAGAAAAGTCAAAGCAGAGTGATGTGATACAGAGTAAT GGTGGGGAGAGGGAGGGCTCCCTGAAGAAGTGAATTGAGAAGCGCATGTCAAG GGTGGGGAGAGGGAGGGCTCCCTGAAGAAGTGAATTGAGAAGCGCATGTCAAG
36481	GGTAGCCACAGTGCCTGGCTTAGGAAAGACTGGTTCTAGGAAAACAATTTCATTCCCTG TGGCCAGCTCAAGCCTCCCCGCCAACGCTTCTCATTCAAGGTCTCTGTGAATTAA AATTCACTCCATCCATCAAACAAGTATTACTGAGCACTAATATGTGCTAGGTACTGCTCC AGGTGCTGAGGAAGTCAGCAGTGAAAGATGACTGCTACTCTCATGGGACATACAGGATAG TAGGAAAAGACAGATAATCAACAAGGTCAATTCTGACCACATCTGTGGTTAAGAAAAA [G, A] TCAAGCAGAGTGATGTGATACAGAGTAATGGTGGGGAGAGGGAGGGCCTCCCTGAAGAAG TGACAGTGAAATTGAGAAGCGCATGTCAAGGGGTTGCCAGGCAGAGGAAATAGGACCCACA TGGGCCTAGAGTCAGGGTGAAGTGCTGAGGAACCTAAAGGCCAATGTGACCA GAGGGAAAGTGAAACAAGGTAAAAAGTGGCAGGGGCCAGGTCCCTAGATGCTTCTAAGC AGTAGAGTGATATGCTCTGGCTACCCCTGGGTCCGTGACCTGGACTGGAAGAAAGCA
36619	ACAAGTATTACTGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGAAGTCAGC AGTAAAAGATGACTGCTACTCTCATGGGACATACAGGATAGTAGGGAAAAGACAGATAA TCAACAAGGTCAATTCTGACCACATCTGTGGTTAAGAAAAGTCAGCAGAGTGATGTG ATACAGAGTAATGGTGGGGAGAGGGAGGGCCCTCCCTGAAGAAGTGACAGTGAAATTGAGAA GCCATGTCAAGGGTTGCCAGGCAGAGGAAATAGGACCCACATGGGCCAGAGTCAGGA [G, A] TGAGCTTGAACTGCTCTGAGGAACCTAAAGGCCAATGTGACCAAGGGAAAGTGAAACAAGGT GAAAAGTTGGCAGGGGCCAGGTCCCTAGATGCTTCTAAGCAGTAGAGTGATATGCTCT GGCTTACCCCTGGTCCGTGACCTGGACTGGAAGAAGCAAGGGTGGACCTGGAAAGA CCACTAGGAGGCTGCTGTTGATGGGTGAGAGAGGAAGGGGGCTGAGAGTAGGGTCAGGGC AGAGGGAGGAGAGCCTGCTGGCTGGCGATGGATGATGGGAAGAGGAACAAAGGA
37088	GACCTGGAAAAGACCACTAGGAGGCTGCTGTTGATGGGTGAGAGAGGAAGGGGGCTGAGAG TAGGGTCAGGGCAGAGGGAGAGACGCTGTCGTTGGCTGGCGATGGATGATGGGAAG AGGAACAAAGGATGACTTTGGTTGGGTCTAAGAAACTGGTGGATGATTGAGCAGG TAGAGAAAAAAATCAGCGTGGGAGGAAAAAAATCAAGACTCTGTTGGACATGGTCA AACTGCCTTCAGACATCCACATAGAGGTATCAGGATACAGAAGTTGGAACACAGAG [G, C]

0 0 0 0 0 0 0 0 0 0

	AAGTCAAGGCTGGAGATTGAAAAAAAAAAAAAAAGTGGGTTATTAGCATA GAGGGCCAATATGGTGAACCCCTGTCTACTGAAAATACAAAATTATCCAGGCATGGT GGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCTGAGGCAGGAGAATTGCTTGAACCC AGAGATGGGTGGAGGTTGCAGTAAGCTGAGATCGGCCATTGCACTCCAGCCTGGGTGA CAGGGCAAGATTCCATCTAAAAAAAGCCACTACAGGATCAACTAAGAGCTCCTA
37204	GAAGAGGAACAAAGGATGACTTTGGTTGGGTCTAACAAAAGTGGGTGGATGATTGAG CAGGTAGAGAAAAATCAGCGTGGGAGGAAAAAAATCAAGACTTCTGTTTGGACATGG TGCAAACCTGCCTTCCAGACATCCACATAGAGGTATCAGGATACAGAAGTTGGAACCTCAC AGAGGAAGTCAGGCTGGAGATTGAAAAAAAGTGGGTATTAGCAGGATCAACTAAGAGCT [G, A] TGGTGGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCTGAGGCAGGAGAATTGCTTGA ACCCAGAGATGGGTGGAGGTTGCAGTAAGCTGAGATCGGCCATTGCACTCCAGCCTGG GTGACAGGGCAAGATTCCATCTAAAAAAAGCCACTACAGGATCAACTAAGAGCT CCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGGCAACTACCTAGCCCTGGGCAT TCATTCCAGCTTCAACTCCAGTGGATGAGGTTATGAGGAGATGGCTTGTGAGGAGATGGAA
37485	ATACAAAATTATCCAGGCATGGTGGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCT GAGGCAGGAGAATTGCTTGAACCCAGAGATGGGTGGAGGTTGCAGTAAGCTGAGATCGT GCCATTGCACTCCAGCCTGGGTGACAGGGCAAGATCCATCTAAAAAAAGCCAC TACAGGATCAACTAAGAGCTCCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGGCC AACTACCTAGCCCTGGGCATTCCAGCTTCAACTCCAGTGGCTTGTGAGGAGATGAGAAGGAGA [G, A] TGTGGAGGTAGATGGAAATGAGAAACAATGCTGTGTCAGAGAGCTAACAGAGTCAGT GTTCAAGAGAGACAGAGCTGTCAACTTGTGATGGATGCTTGTGAGAAGGCAAGCAAGTTG AAGACAAAAAAATGATCTTGGCTCTGCCATATGGCGATCGTGGTGGCCAG GGCAGAGCTTCCATCCAGCGATGGAGACTGCAGACTGGCTGGAGCAGCAGAGAGA AGGAGAGATTAGGAAGTGTGCCAGCACCTATAGACAGCTTCCAGAAGTTATGAGAA
37624	GGTGACAGGGCAAGATTCCATCTAAAAAAAGCCACTACAGGATCAACTAAGAGC TCCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGGCAACTACCTAGCCCTGGGC TTCAAGGAGCTTCAACTCCAGTGGAGAGATGAGAAGGAGAGTGTGGAGGTTAGATGGGAA ATGAGAAACAATGCTGTGTCAGAGAGCTAACAGAGTCAGTGTGTTCAAGAGAGACAGAG CTGCAACTTGTGATGGATGCTTGTGAGAAGGCAAGCAAGTTGAAGACAAAAAA [-, A] TGATCTTGGCTCTGCCATATGGCGATCGTGGTGGCCAGGGCCAGAGCTTCCATCCAG CGATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGC TGCCAGCACCTATAGACAGCTCTCCAGAAGTTATGAGAAGTAACAGCCACGGTCACTG GAGGGACATGGATCAAAGAAAGGGCAGGTGAAGGAGGGAGATGTCGGAGCAGGTTGT TACTGACGAGAAGGAACCACTAGAAAGGGAGAAACTGATGCACTCATCAAACCCCTGTAA
37685	CCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGGCAACTACCTAGCCCTGGGC TCATTCCAGCTTCAACTCCAGTGGAGAGATGAGAAGGAGAGTGTGGAGGTTAGATGGGAA TGAGAAACAATGCTGTGTCAGAGAGCTAACAGAGTCAGTGTGTTCAAGAGAGACAGAGC TGTCAACTTGTGATGGATGCTTGTGAGAAGGCAAGAACAGTGTGAGACAAAAAA TGATCTTGGCTCTGCCATATGGCGATCGTGGTGGCCAGGGCCAGAGCTTCCATCCAG [C, T] GATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGC GCCAGCACCTATAGACAGCTCTCCAGAAGTTATGAGAAGTAACAGCCACGGTCACTGG AGGGGACATGGATCAAAGAAAGGGCAGGTGAAGGAGGGAGATGTCGGAGCAGGTTGT ACTGACGAGAAGGAACCACTAGAAAGGGAGAAACTGATGCACTCATCAAACCCCTGTAA CACGATCATCTCTGTGAATTAGTGTGGTCTGGGAATAGCATCGGAATCAGCCG
37769	AGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAAATGAGAAACAATGCTGTGTCAGAG AGCTAACAGAGAAACTGAGTGTGTTCAAGAGAGACAGAGCTGTCAACTTGTGATGGATGCTTGT AGAAGGCAAGCAAGTTGAAGACAAAAAAATGATCTTGGCTCTGCCATATGG CGATCGTGGTGGCCAGGGCCAGAGCTTCCATCCAGCGATGGAGACTGCAGACTGGCTGG AGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGTGCCAGCACCTATAGACAGCTCTT [C, T] CCAGAAGTTATGAGAAGTAACAGCCACGGTCACTGGAGGGAGATGGATCAAAGAAAGGG CAGGTGAAGGAGGGAGATGTCGGAGCAGGTTGTACTGACGAGAAGGAACCACTAGAA AGGGAGAAACTGATGCACTCATCAAACCCCTGTAAATCACGATCATCTCTGTGAATTAA

FIGURE 3, page 29 of 33

DRAFT - 020004

	GTTCTGGTCTGGAAATAGCATGGGAATCAGCCGCGTGACCTTAGCATTATTCTG TCACTGTTACGATAGACTTGAGTTCTCAGTTCTAAGAAAGTGGAAATAACTACCT
38897	AGAGGGTCTGGATCAAAGGTATTTACACCCAGGGATATTCAGATAAAATCTTTCATC TATGTGAAAACATACAAAGTGGCGAAGTGAGAAACTCCGATTCCTAAGGGTGACAAG TCAAGTGCAGTAATGATGTCATGGTAACCAATATGTTCCAAACTTCCTAAGGGTGACT AGCCCCATGCACTTGGAGAAGTTGGAAATAGGATGTCGTCGTTTATAAAATTGAAAA CACGGTGTCTGCAATCACAGCCACTCACAAAGGAAGCCAGAGATGGTCCCAGCCCCCTCC [G, A] CAGACTTCCTGTGGACTCAGGACTGGGGCTCTCCGGGCTTGCTGTACCCGGAAAT CCAGGGGACAGACTCAGGGTCTGCCCTGCGACAGATGTCGCTAGCCTCTGTGTG CATAAGTCAACTCCGCTCAGGCCAGGGCTGCTGGGCTCCCTGCTGTGGGCAAAACAG CCACCTCGCTGTTCTATCCCCAACCCGTTCCCAGGGAGGGGGCTGGGTGAGACA CCCCCTCAGAGAGGAAGTGTCTCCAGCTTGGAGAGAATCAGGGTGTCTTTCTCT [A, G]
40155	AGACTGAGTGACAGAGTGAGACTGTCTAAAAATAAAAGTGCATCAAGCAGCTGCCCCGTG CCAGGCAGTATACTAGGATCTGGGATCGGGAGGCAAAGATAAAATAGACTCAGTGTCTG TTCTGGAGCCTGCAATGGTCTTCCTCCCTCGCCACACCCACTGCCCTGCGCTGGCCAC CTTCGAAGCCTGTGACTTGCTCCCCAGCTCTCCCTCCCTCTCCATCCACCCCTACA CTTGGCTGCCAGACACAGATAGACCTCTGGAAATAACTGCCCATCAAGGCTGCTTG [A, G] ATCCTGCCTGATCCCTACTGCCATTGACCAGAGTCTGGAGGGAGGGTCAACCTCCCTCC ATGATAACACACTGCACTCTGGCCGGTGGATCCATCTCCAGGAAGCCCCACGACTGCC GCATCCAGGCCCTTCTTGCATCTGCTGGAGGTTCATCTCCATCTGCTATGAG AACATCCGCCTCCCTCCAGGTCCAGATGTTGCCTTACTAAGCGATGGTTCACCGTCTC TTACCTACCATTGTCAGACACTGACCCATGTGGGTCTCTTTCTATTGTACC
40355	CTCCCCAGCTCCTCTCCCTTCTCCATCCACCCACTGCTGCCAGACACAGATA GACCTCCTGAAATAACTGCCCCATCAAGGCTGCTTGAATCCTGCGTATCCCTAC TGCCCATTGACCAGAGTCTGGAGGGAGGGTCACTCCCTCCATGATAACACACTGCACTCC TGGCCGGTGGATCCATCTCCAGGAAGCCCCACGACTGCCGATCCAGGCCCTTCCAG [G, C] TCCAGATGTTGCCCTTACTAAGCGATGGTTCACCGTCTCTAACCTACCATTCTGTCTC CAGACACTGACCCATGTGGTCTCTTCTATTGTACCTCTCATGAGACACCGACCCA GTCTCCTTATGATGATGTTCTGCACATCTCAACTCCTGGGCCACAAGAAAA GATGTCACATCTAACCTCCAGTCTCATCACAGCTCCAGCAAGGGGGCTAAACACAGC ACGTGCCAATTCACTCACTGAGAGGGAGGTGGAGAGGGCATAGGAAGGCAAGAACG
40486	CAGAGTCTGGAGGGAGGGTCACTCCCTCCATGATAACACACTGCACTCCTGGCGGTGGA TCCATCTCCAGGAAGCCCCACGACTGCCGATCCAGGCCCTTCTTGCATCTGTT CCTGGAGGTTCATCTCCATCTGCTATGAGAACATCCGCCTCCCTCCAGGTCCAGATGTT GCCTTACTAACGATGGTTCACCGTCTCTAACCTACCATTCTGTCTCCAGACACTGA CCCATGTGGTCTCTTTCTATTGTACCTCTCATGAGACACCGACCCAGTCTCTTTA [T, C] GATGTGATTGTTCTGCACATCTCAACTCCTCTGGCCACAAGAAAAGATGTCACATC TTAACCTCCAGTCTCATCACAGCTCCAGCAAGGGGCTAACACACAGCACGTGCCAAT TCACATTCACTGAGAGGGAGGTGGAGAGGGCATAGGAAGGCAAGAACGACACGATCTG CCACATGCCCTCCCTCCGGCCCTCTGATTTGGGATCTTCTACTACAAACCCA GCTGCTTCCATGTCGCCCTCCCTGATTTCTGGTAGTCTGGATGGGAGAATGGGG
40512	CTCCATGATAACACACTGCACTCCTGGCGGTGGATCCATCTCCAGGAAGCCCCACGACT GCCGCATCCAGGCCCTTCTTGCATCTGTTCTGGAGGTTCATCTCCATCTGCTA TGAGAACATCCGCCTCCCTCCAGGTCCAGATGTTGCCTTACTAACGATGGTTCACCG TCTCTTACCTACATTCTGTCAGACACTGACCCATGTGGTCTCTTTCTATTG TACCTCTCATGAGACACCGACCCAGTCTCTTATGATGATGATTGTTCTGCACATCTCA [A, C] CTTCCCTCTGGGCCACAAGAAAAGATGTCACATCTAACCTCCAGTCTCATCACAGCTT CCAGCAAGGGGCTAACACACAGCACGTGCCAATTCAACTGAGAGGGAGGTGGAG AGGGCATAGGAAGGCAAGAACGACACGATCTGCCACATGCCCTCCCTCCGGCCCTT CTGATTTGGGATCTTCTACTACAAACAGCTGTCCTCCATGTCGCCCTCCCT GATTCTGGTAGTCTGGATGGGAGAATGGGGACAGTTGTGACCACGAGGAAGCAGAG

FIGURE 3, page 30 of 33

40622	CCATCTGCTATGAGAACATCCGCCCTCCCTCCAGGTCCAGATGTTGCCCTTAACCGATGGGGCTCCTGCACATCTCAACTCCCTGGGCTACAGGAAAGATGTCACATCTAACCTCCAGTCTCATCACAGCTTCCAGCAAGGGGCTAAACACAGCACGTGCCAATTACATTCACTGAGA[-, A, G] GAGAGTGGAGAGGGGCTAGGAAGGCAAGAACGACACGATCTGCCACATGCCCTCCCTCCCAGGAGACACCAGCTCCCTTATGATGTTGATGTTCTGCACATCTCAACTCCCTGGGCTACAAGGGACAGGGCTCTGCCATTGGTACCTATCA GTTCCAATCTTCAAATCAGGTTGATGCCAAGGAAACGCTGGTGGAGAAACCAAAAG
40654	GGTCCAGATGTTGCCCTTAACCGATGTTCACCGTCTTACCTACCATCCCTGTC TCCAGACACTGACCCATGTGGTCTCCTTTCTATTGTCACCTCTCATGAGACACCGACC CAGTCCTCTTATGATGTTGATGTTCTGCACATCTCAACTCCCTGGGCCACAAGAA AAGATGTCACATCTAACCTCCAGTCTCATCACAGCTTCCAGCAAGGGGCTAAACACA GCACGTGCCAATTACATTCACTGAGAGGAGAGTGGAGAGGGCATAGGAAGGCAAGAA [T, C] GCACACGATCTGCCACATGCCCTCCGGCCCTCTGATTTGGGATCTTCATCT ACTACAAAACCAGCTGCCCTCCATGTCGCCCTCCCTGATTTCTGGTAGTCCTGGGAT GGGAGAATGGGACAGTGTGACCACGAGGAAGCAGAGTGGGAGTTCTACAGGCCAAC AGGGCTCTGCCATTGGTACCTATCAGTCCAACTTTCAAATCAGGTTGATGCC CAAGGAAACGCTGGTGGAGAAACCAAAAGAAGGTTCTAGCTGGGTGTTGACCTCTTAGAG [C, G]
40933	AGGGGCATAGGAAGGCAAGAACGCCACACGATCTGCCACATGCCCTCCCTCCCCGGCCCT CTGATTTGGGATCTTCATCTACTACAAAACCAGCTGCCCTCATGTCGCCCTCCCT GATTTCTGGTAGTCCTGGGATGGGAGAATGGGACAGTTGTCACGCCAGAGGAGCAGAG GTGGGAGTTCTACAGGCCAACAGGGCTCTGCCATTGGTACCTATCAGTCCCAATC TTTCAAATCAGGTTGATGCCAAGGAAACGCTGGTGGAGAAACCAAAAGAAGGTTCTAG [C, G] TGGGTGTTGACCTCTTAGAGGCCCATCCGCTAAAGAGGGTTGGGACAGCCTAAATG AGGGAGCTTACAAAAGGGAGCTCTGTGAAACAGTCAGGGTTATCGCAGCATCTCAGG AAATGGGACTAGGCAAGTCTGGCTGGTAGGGATGGTAGGTTACGGAGATCCTTCCACTG ACCCCGCTCTCCACAGGAGAGGCCCTACCTCTACACCCCTGCCAACAGCCGATTG CTCCCAGACCAGGCTGAAGAACGCCCGAGGTGCTACTGGCTGTCCCAGTGGCAGCG
41171	TCTTCAAATCAGGTTGATGCCAAGGAAACGCTGGTGGAGAAACCAAAAGAAGGTTCT AGCTGGGTGTTGACCTCTTAGAGGCCCATCCGCTAAAGAGGGTTGGGACAGCCTAA ATGAGGGAGCTTACAAAAGGGAGCTCTGTGAAACAGTCAGGGTTATCGCAGCATCTC AGGAATGGGACTAGGCAAGTCTGGCTGGTAGGGATGGTAGGTTACGGAGATCCTTCCA CTGACCCCGCTCTCCACAGGAGAGGCCCTACCTCTACACCCCTGCCAACAGCCG [T, C] TGCTCCCAGACCAGGCTGAAGAACGCCCGAGGTGCTACTGGCTGTCCCAGTGGCAG CGGGAGTCGGCCTGTGGTAGGCATCATCGGACAGTGGCTGGAGGTGCAAGGAGAGTG GCCGGAGGAAGATCCTCAAATGAGCCTGCCACTGTGGCCCTCTAGGCCGGGGTGG TCCTCACCCCTAAACTCCCTATAGCCACTCTTCACCCGCCCTCTGCCAGCCACTCCCG GCCACCCAGGACATGCTTCAATGCCCTGGTAGGAAGGCCAGGGGACAGTCACAACTT
41379	TGGTGATGGATGGTCACGGAGATCTTCCACTGACCCCGCTCTCCACAGGAGA GCCCTTACCTCTACACCCCTGCCAACAGCCGATTGCTCCAGGCCAGGCTGAAGAACGCC CCGAGGTGCTCTACTGGCTGTCCCAGTGGCAGCGGGAGTCGGCCTGTGGTAGGCATCA TCGGGACAGTGGCTGGAGGTGCAAGGAGAGGTGGCCGGAGGAAGATCCTCAAATGAGCC TGCCTGACTGTGGCCCTCTAGGCCGGGGGGTGGCTCACCCCTAAACTCCCTATAGCCAC [T, C] CTCTTACCGCCCTCTGCCAGGCCACTCCGCCACCAGGACATGCTTCAATGCCCTGGT GCCATAGGAAGGCCAGGGACAGTCACAACCTCTTGGGCCCTGGCTCTCCAGGCC TGCTCTGGATGGACATCTTAAATAAGCTTAAAGCTTAAACTCCCTAAACTCCCTATAGCC GCGTGAAGCACTACCTTCCATCTTGTGCAAGGCCAGGTGAGGAGCTGCCACTTTTG TGGCTGCCCTCCAGCAGGGCTGCCAAGGCCACGCCAACAGGCCAAACTGCCCTGCCA
41388	ATGGTTCACGGAGATCTTCCACTGACCCCGCTCTCCACAGGAGAGGCCCTTAC TCTACACCCCTGCCAACAGCCGATTGCTCCAGGCCAGGCTGAAGAACGCCGGAGGTGC TCTACTGGGCTGTCCCAGTGGCAGCGGGAGTCGGCCTGTGGTAGGCCACTTTTG TCTACTGGGCTGTCCCAGTGGCAGCGGGAGTCGGCCTGTGGTAGGCCATCATCGGGACAG

	TGGTCTGGAGGTGCAGGAGAGGTGGCCGGAGGAAGATCCTCCAATGAGCCTGCGCACTG TGGCCCTCTAGGCCCGGGGTGGTCCTCACCTAAACTCCATAGCCACTCTCTTAC [C, T]
	GCCCTCTGCCAGGCCACTCCGGCACCAGGACATGCTCAATGCCCTGGGCCATAGGA AGCCCAGGGACAGTCACAACCTCTGGGCTGGCTCTCCAGGCCATGCTCTGG AATGGATACTTAAATAAAAGTCAAAGCTATTCTCTGGGTTGCCCTGCGTGAAGC ACTCACCTCCATCTTGCGAGCCAGGTGTGGAGCTGCCACTTTGTGGCCTGCC TCCAGCAGGGCTGCCAAGCCACGACCAACCAGAGCCAAACTGCCTGCCACCACGAGCA
41880	ATCTCTTGTGCAGGCCAGGTGTGGGAGCTGCCACTTTGTGCCCTGCCCTCAGCAGGGC TGCCCAAGGCCACGACCAACCAGAGCCAAACTGCCTGCCACAGAGCATATCCTCAAGT CACCAAACCCACTATTCTAAAGGCAGAAAAAATGCTGGTCACCAGGTGGCTGGAAATT TTGGAGCTGGCTGGTTGCCATTAGTCAATCCAAACACATACTTAAAGCAACTGTTT GTATCCAGGACAATGCGAACACTGAGGTGCCCTAGGCTGTGCATGTCGAGCCTGGC [A, C] GAGAGGTCAAACCTTCATAACCAAGAACGCCACGTGATGATGTGTAACACTAGGGCA TCAGTAGGTAATGTGCTGATTGTTAAAGAATAGAAAGGGTCTCGGGGAAAGTTT CTTGGGGAGAGCACCTTCACATGTCATTGGGAAAGAATAAAAATGATTGGGAC ACAAATACCTCTATATTCTCAACCTGATTCTCAAGGTGCTAAATTAGGAAAAATT CCTATTCTATATGCCAGGTTCTGAGGGAAAAGTAGAGAGTCTGAAAATATGGCCT
42278	AAAGGGTTCTCGGGGAAAGTTCTGGGGAGAGCAACCTTCACATGTCATTGGGAA AAGGAATAAAAATGATTGGACACAAATACCTCTATATTCTCAACCTGATTCTCAA GGTGTCAAATTCTACGAAAAAATTCCTATTCTATGCCCAGGTTCTGAGGGAAAACTA GAGAGAGTCTGAAAATATGGGCTGCATTCACTGAGCCCCCTGCTAGGGCGAGGCCCCGTG CTGGAGGCCTTCCACAGATGGTCTCTTTATGCTGCACAAAAGCCAGGGAGGGGGTAAA [G, A] GGAAAATCTTGAAAATAGAAGTGATGCTTGCACACCGTGAATGTAACACGCCGC GAATTGTTCCATTAAATGATTAATTGTGATCATGTGAATTCACTTCATAAAAG AATCCAGGGAGGGTAGACATCATCTGCAATTGTAACCTCTCTGATCCTGAAGTCCGGGA TGATAAAGAGCCTGAGTCACATCCGGATGCAACACTGAAATGCTGTGCCCTGAAGCTG CCTCGCCAGCCTGAGCCAGTGTCCAGGCTCTGCATCTGTAAGGAGTAAAGAGT
42339	AGGAATAAAAATGATTGGGACACAAATACCTCTATATTCTCAACCTGATTCTCAAG GTGCTAAATTCTAGGAAAAAATTCCTATTCTATATGCCCAGGTTCTGAGGGAAAACTAG AGAGAGTCTGAAAATATGGGCTGCATTCACTGAGCCCCCTGCTAGGGCGAGGCCCCGTG TGGAGGCCTTCCACAGATGGTCTCTTTATGCTGCACAAAAGCCAGGGAGGGGGTAAA GGAAAATCTTGAAAATAGAAGTGATGCTTGCACACCCGTGAATGTAACACGCCGC [G, A] AATTGTTCCATTAAATGATTAATTGTGATCATGTGAATTCACTTCATAAAAAGA ATCCAGGGAGGGTAGACATCATCTGCAATTGTAACCTCTCTGATCCTGAAGTCCGGGAT GATAAAGAGCCTGAGTCACAAATCCGGATGCAACACTGAAATGCTGTGCCCTGAAGCTG CTTCGCCAGCCTGAGCCAGTGTCCAGGCTCTGCATCTGTAAGGAGTAAAGAGT CACATTGCTTATCTCACGGCCTGCTGAAAGAACCGTGTGAACCTCTAAC
42612	CAACACCGTGAATGTAACACGCCGAATTGTCATTAAAATGATTAATTGTGAT CATGTGAATTCACTTCATAAAAAGAATCCAGGGAGGTAGACATCATCTGATTGTA ACCTCTCTGATCCTGAAGTCCGGGATGATAAAAGAGCCTGAGTCACAAATCCGGATGCA ACACTGAAATGCTGTGCCCTGAAGCTGCCCTGCCAGCCTGAGCCAGTGTCCCAGGCTC TGCATCTGAAAATGGAGTAAGAGTACACATTGCTTATCTCACGGCCTGCTGAAA [A, G] ATAAGGAACCGTGTGTGAACCTCTAACCTAAATGCTGCACAACTGAAAATGCCCTTT TCCTCGGTGAAGAGTTGGGATAAGGCCAGACTGTTGGGAAAGATGTGAGACCCAGAGAT GAGTTGGGAAATGGGTAAATAACATATGGGTGAGAGTGCCTGCCCTCTCAGGGGA GGTCATCACCTTATCTCTGTCACAACAGAGAACCCGGAGGACCTATACCCAGTTC CGTGTCTTCTGGCTCAGTGTCTGTTCTATACAATGGGAACAGCATGCATTCCCCCTG
42817	TGCCCTCGCCAGCCTGAGCCAGTGTCCAGGCTCTGCATCTGAAAATGGAGTAAGA GTACACATTGCTTATCTCACGGCCTGCTGAAAATAAGGAACCGTGTGTGAACCTCT AACTCTAAAATGCTGCACAACGAAATGCCCTTTCTCGGTGAAGAGTTGGGATAAG GCCAGACTGTTGGGAAAGATGTGAGACCCAGAGATGAGTTGGGAAATGGGTAAATAA CATATGGGTGGAGAGTGCCTCTCAGGGAGGTTCATCACCTTATCTCTTCTG [T, G]

FIGURE 3, page 32 of 33

CACAAACAGAGAACCGGAGGACCTATACCCAGTTCCGTGTTCTGGGCTTCAGTGTCT
GTTTCTATAACATGGGAACAGCATGCATTCCCCTGCTTTCTATAGACTGGAAAACGT
GGTGACCAAGTCACACATCCCAGCTTATGCTCCCGCTTAAGACAGTGTAAACGACAAAGG
TAACCCTTACACTCCTGGTTTGAGACAGTATAACGACAAAGGTAAACATAGGAAGTCAAGG
AGTCGCTTCACCGCCCCCTCCCCCACCCTTTTTCTGCAAGTTCTATT